

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2005, 04:10:34 ; Search time 160 Seconds
(without alignments)
6381.479 Million cell updates/sec

Title: US-10-618-408-3
Perfect score: 624
Sequence: 1 cgcgcgctntaagcattt.....tccacagactgctgcgcgnaca 624

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	566.2	90.7	2052	3	US-08-630-916A-45
C 2	566.2	90.7	2052	3	Sequence 45, Appl
C 3	246.6	39.5	2559	2	US-09-657-481A-3
C 4	246.6	39.5	2559	3	Sequence 3, Appl
C 5	246.6	39.5	2559	3	US-09-070-060-2
C 6	246.6	39.5	2559	3	Sequence 2, Appl
C 7	246.6	39.5	2559	3	US-09-357-746-2
C 8	246.6	39.5	2559	3	Sequence 234, App
C 9	246.6	39.5	2559	3	Sequence 1, Appl
C 10	246.6	39.5	2559	3	US-09-357-746-1
C 11	246.6	39.5	2559	3	Sequence 1, Appl
C 12	246.6	39.5	2559	3	US-09-070-060-1
C 13	246.6	39.5	2559	3	Sequence 10, Appl
C 14	246.6	39.5	2559	3	US-09-657-481A-10
C 15	246.6	39.5	2559	3	Sequence 47, Appl
C 16	246.6	39.5	2559	3	US-08-630-916A-47
C 17	246.6	39.5	2559	3	Sequence 100, App
C 18	246.6	39.5	2559	3	US-09-774-639-100
C 19	246.6	39.5	2559	3	Sequence 2, Appl
C 20	246.6	39.5	2559	3	US-08-895-601-2
C 21	246.6	39.5	2559	3	Sequence 1035, Ap
C 22	246.6	39.5	2559	3	US-09-949-016-1035
C 23	246.6	39.5	2559	3	Sequence 3, Appl
C 24	246.6	39.5	2559	3	US-08-539-205A-3
C 25	246.6	39.5	2559	3	Sequence 3, Appl
C 26	246.6	39.5	2559	3	US-09-392-163A-3
C 27	246.6	39.5	2559	3	Sequence 4083, Ap
C 28	246.6	39.5	2559	3	US-08-539-205A-5
C 29	246.6	39.5	2559	3	Sequence 5, Appl
C 30	246.6	39.5	2559	3	US-09-392-163A-5
C 31	246.6	39.5	2559	3	Sequence 239, App
C 32	246.6	39.5	2559	3	US-09-401-064-239
C 33	246.6	39.5	2559	3	Sequence 714, App
C 34	246.6	39.5	2559	3	US-09-799-451-714
C 35	246.6	39.5	2559	3	Sequence 4084, App
C 36	246.6	39.5	2559	3	US-09-248-796A-4084
C 37	246.6	39.5	2559	3	Sequence 49, Appl
C 38	246.6	39.5	2559	3	US-08-630-916A-49
C 39	246.6	39.5	2559	3	Sequence 3075, Ap
C 40	246.6	39.5	2559	3	US-09-949-016-3075
C 41	246.6	39.5	2559	3	Sequence 2160, Ap
C 42	246.6	39.5	2559	3	US-09-270-767-2160
C 43	246.6	39.5	2559	3	Sequence 17442, A
C 44	246.6	39.5	2559	3	US-09-270-767-17442
C 45	246.6	39.5	2559	3	Sequence 28116, A
C 46	246.6	39.5	2559	3	US-09-270-767-28116
C 47	246.6	39.5	2559	3	Sequence 12361, A
C 48	246.6	39.5	2559	3	US-09-270-767-12361
C 49	246.6	39.5	2559	3	Sequence 1, Appl
C 50	246.6	39.5	2559	3	US-08-476-509B-1

Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 3, Appli
Sequence 3, Appli
Sequence 14, Appli
Sequence 15085, A
Sequence 226, App
Sequence 1, Appli
Sequence 1, Appli
Sequence 3, Appli
Sequence 12556, A
Sequence 16078, A
Sequence 824, App
Sequence 2436, Ap
Sequence 14178, A
Sequence 5403, Ap

ALIGNMENTS

RESULT 1
US-08-630-916A-45/c
; Sequence 45, Application US/08630916A
; Patent No. 6011137
; GENERAL INFORMATION:
; APPLICANT: Pirozzi, Gregorio
; APPLICANT: Kay, Brian K.
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
; TITLE OF INVENTION: POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,916A
; FILING DATE: 03-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-203
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 896-8864/9741
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2052 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-08-630-916A-45

Query Match 90.7%; Score 566.2; DB 3; Length 2052;
Best Local Similarity 97.3%; Pred. No. 5.3e-172;
Matches 582; Conservative 0; Mismatches 15; Indels 1; Gaps 1;
Qy 27 CCCTTAGTACAGTACTTCCCATCGAGGAGTCTTGAATCTTGTGTTCTTGTGTTA 86
Db 1044 CCTTTAGTACAGTACTTCCCATCGAGGAGTCTTGAATCTTGTGTTCTTGTGTTA 985

QY 87 TGATCAACAAAGTACCTTACACCTTCAACAGTATATCTAATTTCCAGGCTTCTGGCAGG 146

DB 984 TGATCAACAAAGTACCTTACACCTTCAACAGTATATCTAATTTCCAGGCTTCTGGCAGG 925

QY 147 GGTCTCTTCATCTGTAAAGCTTGTAGTCTTCCACCTTCCAGCTTGTGTTTGTGTTA 206

DB 924 GGTCTCTTCATCTGTAAAGCTTGTAGTCTTCCACCTTCCAGCTTGTGTTTGTGTTA 865

QY 207 TGATTCACAAAGTAAACCTCTGTCTGTGTAATCCACTCTTTTTCACAGCTGGTGGCAAA 266

DB 864 TGATTCACAAAGTAAACCTCTGTCTGTGTAATCCACTCTTTTTCACAGCTGGTGGCAAA 805

QY 267 GGTCCATAAAGGTCATTTTCTGCAGTAAACCTTGAATCCACTCTTTTTCACAGCTGGTGGCAAA 326

DB 804 GGTCCATAAAGGTCATTTTCTGCAGTAAACCTTGAATCCACTCTTTTTCACAGCTGGTGGCAAA 745

QY 327 AACTGTTGTCATAGCTTCTGCAGTAAACCTTGAATCCACTCTTTTTCACAGCTGGTGGCAAA 386

DB 744 AACTGTTGTCATAGCTTCTGCAGTAAACCTTGAATCCACTCTTTTTCACAGCTGGTGGCAAA 685

QY 387 ACAGATTCCATAGTGGCGCTGCCACCTGTTGTTTCTGTTGTTATGATCCACATAATAA 446

DB 684 ACAGATTCCATAGTGGCGCTGCCACCTGTTGTTTCTGTTGTTATGATCCACATAATAA 625

QY 447 ACTCTCTACGATCATCAACTCTTCTTCCACCTGGAGGTTAAAGGTTGTGTTCTCTCC 506

DB 624 ACTCTCTACGATCATCAACTCTTCTTCCACCTGGAGGTTAAAGGTTGTGTTCTCTCC 565

QY 507 CATGTGTAGTTCGAGCATTAATGATCCACATAATAAGGTTGTGTTCTCTCTCTCTCC 566

DB 564 CATGTGTAGTTCGAGCATTAATGATCCACATAATAAGGTTGTGTTCTCTCTCTCTCT 505

QY 567 TGTTCACCACTGATGGCAANGTCTTCTGTTGGCACTTCCAGACTGTCGCCGNACA 624

DB 504 TGTTCACCACTGATGGCAANGTCTTCTGTTGGCACTTCCAGACTGTCGCCGNACA 448

RESULT 2

US-09-657-481A-3/c

; Sequence 3, Application US/09657481A

; Patent No. 6258601

; GENERAL INFORMATION:

; APPLICANT: Lex M. Cowser

; APPLICANT: Brett P. Monia

; TITLE OF INVENTION: ANTISENSE MODULATION OF UBIQUITIN PROTEIN LIGASE WWPI AND WW

; FILE REFERENCE: RTS-0087

; CURRENT APPLICATION NUMBER: US/09/657,481A

; CURRENT FILING DATE: 2000-09-07

; NUMBER OF SEQ ID NOS: 93

; SEQ ID NO 3

; LENGTH: 2052

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1) ... (2052)

US-09-657-481A-3

Query Match 90.7%; Score 566.2; DB 3; Length 2052;

Best Local Similarity 97.3%; Pred. No. 5.3e-172;

Matches 582; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

QY 27 CCTTTAGTACAGATGATCTCCCATTCGAGGATCTTTGATGTTGTTGTTGTTGTTA 86

DB 1044 CCTTTAGTACAGATGATCTCCCATTCGAGGATCTTTGATGTTGTTGTTGTTGTTA 985

QY 87 TGATCAACAAAGTACCTTACACCTTCCAGGATATATCTAATTTCCAGGCTTCTGGCAGG 146

DB 984 TGATCAACAAAGTACCTTACACCTTCCAGGATATATCTAATTTCCAGGCTTCTGGCAGG 925

QY 147 GGTCTCTTCATCTGTAAAGCTTGTAGTCTTCCACCTTCCAGCTTGTGTTTGTGTTA 206

DB 924 GGTCTCTTCATCTGTAAAGCTTGTAGTCTTCCAGCTGGTGTGTTTGTGTTA 865

QY 207 TGATTCACAAAGTAAACCTCTGTCTGTGTAATCCACTCTTTTTCACAGCTGGTGGCAAA 266

DB 864 TGATTCACAAAGTAAACCTCTGTCTGTGTAATCCACTCTTTTTCACAGCTGGTGGCAAA 805

QY 267 GGTCCATAAAGGTCATTTTCTGCAGTAAACCTTGAATCCACTCTTTTTCACAGCTGGTGGCAAA 326

DB 804 GGTCCATAAAGGTCATTTTCTGCAGTAAACCTTGAATCCACTCTTTTTCACAGCTGGTGGCAAA 745

QY 327 AACTGTTGTCATAGCTTCTGCAGTAAACCTTGAATCCACTCTTTTTCACAGCTGGTGGCAAA 386

DB 744 AACTGTTGTCATAGCTTCTGCAGTAAACCTTGAATCCACTCTTTTTCACAGCTGGTGGCAAA 685

QY 387 ACAGATTCCATAGTGGCGCTGCCACCTGTTGTTTCTGTTGTTATGATCCACATAATAA 446

DB 684 ACAGATTCCATAGTGGCGCTGCCACCTGTTGTTTCTGTTGTTATGATCCACATAATAA 625

QY 447 ACTCTCTACGATCATCAACTCTTCTTCCACCTGGAGGTTAAAGGTTGTGTTCTCTCC 506

DB 624 ACTCTCTACGATCATCAACTCTTCTTCCACCTGGAGGTTAAAGGTTGTGTTCTCTCC 565

QY 507 CATGTGTAGTTCGAGCATTAATGATCCACATAATAAGGTTGTGTTCTCTCTCTCTCT 566

DB 564 CATGTGTAGTTCGAGCATTAATGATCCACATAATAAGGTTGTGTTCTCTCTCTCTCT 505

QY 567 TGTTCACCACTGATGGCAANGTCTTCTGTTGGCACTTCCAGACTGTCGCCGNACA 624

DB 504 TGTTCACCACTGATGGCAANGTCTTCTGTTGGCACTTCCAGACTGTCGCCGNACA 448

RESULT 3

US-09-070-060-2/c

; Sequence 2, Application US/09070060

; Patent No. 5976849

; GENERAL INFORMATION:

; APPLICANT: Hustad, Carolyn M.

; APPLICANT: Ghildyal, Namit

; TITLE OF INVENTION: Human E3 Ubiquitin Protein

; TITLE OF INVENTION: Ligase

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ZENECA Pharmaceuticals, Inc.

; STREET: 1800 Concord Pike

; CITY: Wilmington

; STATE: DE

; COUNTRY: USA

; ZIP: 19850-5437

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/070,060

; FILING DATE: 30-APR-1998

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/073,839

; FILING DATE: 05-FEB-1998

; ATTORNEY/AGENT INFORMATION:

; NAME: Higgins, Patrick H

; REGISTRATION NUMBER: 39,709

; REFERENCE/DOCKET NUMBER: PHM.70312

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 302.886.4889

; TELEFAX: 302.886.8221

; TELEX:

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2559 base pairs

; TYPE: nucleic acid

```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-070-060-2

Query Match      39.5%; Score 246.6; DB 2; Length 2559;
Best Local Similarity 64.3%; Pred. No. 7.3e-69;
Matches 388; Conservative 0; Mismatches 200; Indels 15; Gaps 1;

Qy 12 TAAGCATTTGTGGACCCCTTTAGTNCAGATGACATTCCTCCATTCGCGAGGATCTTTGAATGTT 71
Db 1418 TAGGCTATCTGAGTCCATTTGCTAGGCGAGATTTCTCTGTGCGGGATCTATATAGGTG 1359

Qy 72 GTTGTCTTGTGTATGATCAAAAGTACCTTACACCTTCACGAGTATATCTAAATTTCC 131
Db 1358 GTAGTCTTCTATTGTGTGTCACAAAATATGGAATTCATCCACTGTGAATCTCATTTCC 1299

Qy 132 CAGCCTTCGCGAGGGGTTCTTCAATCTGTAAAGCCTTGAGTCTTGGATCTTCCCACTGG 191
Db 1298 CAACCTTCAGTAAGGGCTTTTCAATTAATTTGACCTTGACTCTGGGGTCTTCCCATGT 1239

Qy 192 GTTGTCTTGTGTATGATTCACAAAGTAAACCTGTCTGTGAATCCACTCTTTTTTTC 251
Db 1238 GTAATTCGTGTGTGTGTTGACGAAATATACCTGCAATGCTGTCTGTCTCTCTCC 1179

Qy 252 CAGCTGGTGGCAAGGTCCTAATAGGTCATTTTCT-----GCAGCTAAC 296
Db 1178 CATCCAGGTGGCAATGACCAAGAGGATCAAAATCTTTACTTTGTGATGTAGCAAAATAA 1119

Qy 297 ATTGAGCCGAATAGAGGTATCGTTGGTTAAACTGTTTGCATAGCTCCCTGCAATGGTTC 356
Db 1118 TCTTGATTCCTAATAATGAATCTCTGGTTAAACTGCTGCATTTGCTCTTGAAGCTGACTA 1059

Qy 357 CGCTGAGATTCGCACTGNTCAAAATTCGAGACAGATTCCTATGAGGCGCTGCCAGTT 416
Db 1058 CGCTGAGTGGCAATGTTTCACTAGTTCGAGCGGATTCAGTGTGGCTCTGCGCAGTT 999

Qy 417 GTTGTCTTGTGTATGATCCACATAATAAATCTCTACGATCATCAACTCTTCTTTTC 476
Db 998 GTTGTCTTGTGAAATGGTCAACATAATAAATACGTCCTCCATGTTGTCAACCCGCGTTC 939

Qy 477 CAACCTGGAGTAAAGGTTGTGTCTCTCCCATGTGTAGTTCGAGCATTTATGANCACA 536
Db 938 CAGCCAGGAGTAGAGTTCGTCTATCCATGTTGTCTTCTTCTCAACATGATCTACA 879

Qy 537 TAATACGCTCTACCATGAGGATCTTTTCTTTTGTTCACCCCTGATGGCAANGTTCTGT 596
Db 878 TAGTAAACTCGCCGCTGCTGCTCACTCTCTGCTCCCAACAGGTGGCAAGGAGCTTGA 819

Qy 597 GTT 599
Db 818 GTT 816

RESULT 5
US-09-919-039-234/c
; Sequence 234, Application US/09919039
; Patent No. 6727066
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; PRIOR FILING DATE: 2002-09-09
; CURRENT FILING DATE: 2002-09-09
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 234
; LENGTH: 2678
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6727066 1989186CB1
```

US-09-919-039-234									
Query Match									
Best Local Similarity 39.5%; Score 246.6; DB 4; Length 2678;									
Matches 388; Conservative 64.3%; Pred. No. 7.4e-69;									
Mismatches 200; Indels 15; Gaps 1;									
QY	12	TAAGCAATTTGTGGACCCCTTTATGNCAGATGACATCCCATTCGCGAGGATCTTTGAATGTT	71						
DB	1190	TAGGCTATCTGAGGTCCATTTCTAGGGCAGATTTTCTGTGCGGGATCTATATAGGTG	1131						
QY	72	GTGTCTTGTGTATGATCAACAAGTACCTTACACCTTCACGAGTATATCTAATTTCC	131						
DB	1130	GTAGTTCTTCTATTTGTGGTCCAAATATGGAATTCATTCACCTGTGATCTCATTTCC	1071						
QY	132	CAGCCTTCCTGGCAGGGGTTCTTCATTTCTGTAAAGCTTGTGATTCCTTGGATTCCTCCACGTG	191						
DB	1070	CAACCTTCAGGTAAAGGCTTTTCATTTAATGACCTTGACTTCCTGGGTCTTCCCATGT	1011						
QY	192	GTGTGTTTGTGTATGATCAACAAGTAAACCTGTCTGTTGAATCCACTCTTTTTC	251						
DB	1010	GTAAATTCGTGTGTGTGTTGACGAAATATATCTGCCATGTCTGTTCTCTCTCC	951						
QY	252	CAGCCTGTGGCAAGGTCCATTAAGGTCAATTTCT	296						
DB	950	CATCCAGGTGGCAATGACCAAGAGGATCAATTTCTTACTTTGTGATGTAGCAATATA	891						
QY	297	ATTGAAGCCGAATPAGAGGTATCGTTGGTTAAACTGTGTGATGATCTCCCTGCAATTTGGTTC	356						
DB	890	TCCTGATTTCCATAAATGAATCTCTGTTAAACTGTGCAATGCTCTTGAAGCTGACTA	831						
QY	357	CGCTGAGATGCCACTGNTCAAAATTCGGACAGATTCATGATGAGCGCTGCCACGTT	416						
DB	830	CGCTGATGTCGCAATGTTTATAGTTCCTGGACGAGATTCAGTGTGGCTCTGCCACGTT	771						
QY	417	GTGTGTTCTGTGTATGATCCACATAATAAATCTCTACGATCATCACTCTTTTCC	476						
DB	770	GTGTGTTCTGTGAATGGTCAACATAATAAATACGTCCTCATGTTGTCAACCCCGCTTC	711						
QY	477	CAACCTGGAGGTAAAGTTGTGTCTCTCCCATGTGTGATGTCGAGCATATGANCACA	536						
DB	710	CAGCCAGGAGTAAAGTTGTGTCTCTATCCCATGTTTCTTTCTCAACATGATCTACA	651						
QY	537	TAATACCTCTACCATGAGGATCTTTTCTTGTTCCTCCACCCCTGATGGCAANGTTCTGT	596						
DB	650	TAGTAACTCGCCGTGCTGGTCACTCTCTCTCCCAACCAAGGTGGCAAGGAGCTTGA	591						
QY	597	GTT 599							
DB	590	GTT 588							

RESULT 6
US-09-357-746-1/c
; Sequence 1, Application US/09357746
; Patent No. 6087122
; GENERAL INFORMATION:
; APPLICANT: ZENECA Limited
; TITLE OF INVENTION: HUMAN E3 UBIQUITIN PROTEIN LIGASE
; FILE REFERENCE: PHM.70312.NI
; CURRENT APPLICATION NUMBER: US/09/357,746
; CURRENT FILING DATE: 1999-07-21
; EARLIER APPLICATION NUMBER: US No. 6087122 60/073,839
; EARLIER FILING DATE: 1998-02-05
; EARLIER APPLICATION NUMBER: US No. 608712209/070,060
; EARLIER FILING DATE: 1998-04-30
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 5372
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-357-746-1

Query Match
Best Local Similarity 39.5%; Score 246.6; DB 3; Length 5372;
Matches 388; Conservative 64.3%; Pred. No. 1.1e-68;
Mismatches 200; Indels 15; Gaps 1;
QY 12 TAAGCAATTTGTGGACCCCTTTATGNCAGATGACATCCCATTCGCGAGGATCTTTGAATGTT 71
DB 1567 TAGGCTATCTGAGGTCCATTTCTAGGGCAGATTTTCTGTGCGGGATCTATATAGGTG 1508
QY 72 GTGTGTTCTTGTGTATGATCAACAAGTACCTTACACCTTCACGAGTATATCTAATTTCC 131
DB 1507 GTAGTTCTTCTATTTGTGGTCCAAATATGGAATTCATTCACCTGTGATCTCATTTCC 1448
QY 132 CAGCCTTCCTGGCAGGGGTTCTTCATTTCTGTAAAGCTTGTGATTCCTTGGATTCCTCCACGTG 191
DB 1447 CAACCTTCAGGTAAAGGCTTTTCATTTAATGACCTTGACTTCCTGGGTCTTCCCATGT 1388
QY 192 GTGTGTTTGTGTATGATCAACAAGTAAACCTGTCTGTTGAATCCACTCTTTTTC 251
DB 1387 GTAAATTCGTGTGTGTGTTGACGAAATATATCTGCCATGTCTGTTCTCTCTCTCC 1328
QY 252 CAGCCTGTGGCAAGGTCCATTAAGGTCAATTTCT -----GCAGCTAAC 296
DB 1327 CATCCAGGTGGCAATGAGCAAGAGGATCAATTTCTTACTTTGTGATGTAGCAATATA 1268
QY 297 ATTGAAGCCGAATPAGAGGTATCGTTGGTTAAACTGTGTGATGATCTCCCTGCAATTTGGTTC 356
DB 1267 TCCTGATTTCCATAAATGAATCTCTGTTAAACTGTGCAATGCTCTTGAAGCTGACTA 1208
QY 357 CGCTGAGATGCCACTGNTCAAAATTCGGACAGATTCATGATGAGCGCTGCCACGTT 416
DB 1207 CGCTGATGTCGCAATGTTTATAGTTCCTGGACGAGATTCAGTGTGGCTCTGCCACGTT 1148
QY 417 GTGTGTTCTGTGTATGATCCACATAATAAATCTCTACGATCATCACTCTTTTCC 476
DB 1147 GTGTGTTCTGTGAATGGTCAACATAATAAATACGTCCTCATGTTGTCAACCCCGCTTC 1088
QY 477 CAACCTGGAGGTAAAGTTGTGTCTCTCCCATGTGTGATGTCGAGCATATGANCACA 536
DB 1087 CAGCCAGGAGGTAAAGTTGTGTCTCTATCCCATGTTTCTTTCTCAACATGATCTACA 1028
QY 537 TAATACCTCTACCATGAGGATCTTTTCTTGTTCCTCCACCCCTGATGGCAANGTTCTGT 596
DB 1027 TAGTAACTCGCCGTGCTGGTCCACTCTCTCTCCCAACCAAGGTGGCAAGGAGCTTGA 968
QY 597 GTT 599
DB 967 GTT 965

RESULT 7
US-09-070-060-1/c
; Sequence 1, Application US/09070060
; Patent No. 5976849
; GENERAL INFORMATION:
; APPLICANT: Husted, Carolyn M.
; APPLICANT: Ghildyal, Namit
; TITLE OF INVENTION: Human E3 Ubiquitin Protein
; TITLE OF INVENTION: Ligase
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZENECA Pharmaceuticals, Inc.
; STREET: 1800 Concord Pike
; CITY: Wilmington
; STATE: DE
; COUNTRY: USA
; ZIP: 19850-5437
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070,060


```
/ ADDRESSEE: Pennie & Edmonds
/ STREET: 1155 Avenue of the Americas
/ CITY: New York
/ STATE: New York
/ COUNTRY: United States
/ ZIP: 10036-2711
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/630,916A
/ FILING DATE: 03-APR-1996
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: MISROCK, S. LESLIE
/ REGISTRATION NUMBER: 18,872
/ REFERENCE/DOCKET NUMBER: 1101-203
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 790-9090
/ TELEFAX: (212) 896-8864/9741
/ INFORMATION FOR SEQ ID NO: 47:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 3476 bases
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA
US-08-630-916A-47

Query Match          32.6%; Score 203.6; DB 3; Length 3476;
Best Local Similarity 64.0%; Pred. No. 6.3e-55;
Matches 340; Conservative 0; Mismatches 182; Indels 9; Gaps 2;

QY 47 CCCATTGCGAGGATCTTTGAAGTGTGTTGTTGTTGTTATGATCAACAAAGTACCTTAC 106
DB 1542 CCCCGGCGAGGATCTTAAAGTGTGTTGCGGGTATTGTGTCACAAAGTATCGCAC 1483

QY 107 ACCTTCACGAGTATCTAATTTCCAGCCTTCTGCAGGGTCTTTCATTCTGTGAAGC 166
DB 1482 CCGCTCGCTGGTGTATTTCATCTCCCATCTTGGGGGCAAGCTGCTTCCGATCATCCC 1423

QY 167 TTGAGTCTTGGATCTTCCACCTGGGTGTTTGTGTTATGATTCACAAAGTAAACCTT 226
DB 1422 CTGGGTCCGGGGATCTCTCCACCTGGGTCTGCGAGTGTATTGTTCACTGATACACCG 1363

QY 227 GTCTGTTGAATCCACTCTTTTTCAGCCTGTTGGCAAGGTGCCATAAGGGTCAATTTTC 286
DB 1362 TCCATT--GTCTGTCTTTTCTCCCAACCCAGGAGGGGGGCCAGGGGATCATGGTC 1306

QY 287 TGCACTAACATTGAAGCCGAATAGAGGTATCGTTGTTAACTGTTGCATAGTCCCTG 346
DB 1305 AGTCGAAGCACTCCAAACTGGTATAGGAATCTTTGGCTGAAGTGTGCTGATGGCCCCCTG 1246

QY 347 CAATTGGTTCGCTGAGATGCGCACTGNTCAAAATTCGACAGATTCATGTTAGGCGCG 406
DB 1245 GAGCTGATTCGCTGCGACTGCCACTGCTCATAGTTGGCACGTACTCCGGTTCGGAGC 1186

QY 407 CTGCCACGTTGTTGTTCTGGTGTATGATCCACATAATAAATCTCTCTAGATCATCAAC 466
DB 1185 CTGCCAGGTGGTGGTCCGAGTATTGTATCCATAGTAAACCTGCTCGGGGATCTGT 1126

QY 467 TCTTCTTCCCACTGGAGGTAAGGTGTGTCTCTCCATGTGGTAGTTCGAGCAT 526
DB 1125 GCGTTTTCACGCTGGAGGAA-----GGGGCCGCTCCAGGTGGTGTCTTGGTATT 1072

QY 527 ATGANCACATAATACGCTCTACCATGAGGATCTTTTCTTTGTTCCCAACC 577
DB 1071 GTGGTCAACATATAGACAGTCCGTTGGGACGCTCTCGCTGTTCCTCCATCC 1021

, RESULT 10
```

```
US-09-774-639-100/c
/ Sequence 100, Application US/09774639
/ Patent No. 6806351
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: 90 Human Secreted Proteins
/ FILE REFERENCE: PZ013P1
/ CURRENT APPLICATION NUMBER: US/09/774,639
/ CURRENT FILING DATE: 2001-07-09
/ PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/244,112
/ PRIOR FILING DATE: EARLIER FILING DATE: 1999-02-04
/ NUMBER OF SEQ ID NOS: 371
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 100
/ LENGTH: 2351
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (593)
/ OTHER INFORMATION: n equals a,t,g, or c
US-09-774-639-100

Query Match          17.9%; Score 111.8; DB 4; Length 2351;
Best Local Similarity 64.2%; Pred. No. 2.1e-25;
Matches 167; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 12 TAAGCATTTGTGAGCCCTTTAGTNCAGATGATCCCATTCGAGGATCTTTGAATGTT 71
DB 284 TAGGCTATCTGAGTCCATTGTCTAGGCGAGATTTTCCTGTGGGGATCTATATAGGTG 225

QY 72 GTGTGTTCTTGTATGATCAACAAAGTACCTTACACCTTCACGAGTATATCTAATTTCC 131
DB 224 GTAGTCTTCTATTGTGGTCCACAAATATGAAATTCACCTGTGAATCTCAATTTCC 165

QY 132 CAGCCTTCTGCGAGGGGTCTTTCATTCTCTAAGCCTTGAGTCTTGGATCTTCCCACTGG 191
DB 164 CAACCTTCAGGTAAGGGCTTTTCATTTAATTGACCTTGACTTCTGGGGTCTTCCCAATGT 105

QY 192 GTGTGTTTGTGTATGATTCACAAAGTAAACCTGCTGTGTTGAATCCACTCTTTTTC 251
DB 104 GTAATTCGTGTGTGTGTGTGAGAAATATATCTGCAATGCTGTCTGTCTCTCTCTCC 45

QY 252 CAGCCTGTTGGCAAGGTCC 271
DB 44 CATCCAGTGGCAATGGACC 25

RESULT 11
US-08-895-601-2/c
/ Sequence 2, Application US/08895601
/ Patent No. 6060262
/ GENERAL INFORMATION:
/ APPLICANT: Beer-Romero, Peggy
/ APPLICANT: Strack, Peter J.
/ APPLICANT: Glass, Susan J.
/ APPLICANT: Rolfe, Mark
/ TITLE OF INVENTION: REGULATION OF KAPPA B (Ikb) DEGRADATION,
/ TITLE OF INVENTION: AND METHODS AND REAGENTS RELATED THERETO
/ NUMBER OF SEQUENCES: 16
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: FOLEY, HOAG & ELIOT LLP
/ STREET: One Post Office Square
/ CITY: Boston
/ STATE: MA
/ COUNTRY: USA
/ ZIP: 02109-2170
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/08/895,601
; FILING DATE: 16-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MIV-096.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2790 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..2782
; US-08-895-601-2

Query Match      8.7%; Score 54; DB 3; Length 2790;
Best Local Similarity 48.5%; Pred. No. 1e-06;
Matches 178; Conservative 0; Mismatches 183; Indels 6; Gaps 1;

QY 173 TCTGGATCTCCCACTGGGTTGTTTGTATGATTCACAAAGTAAACCTCTGTCTGT 232
DB 1441 TCITGGATCTCCCACTGGGTTGTTTGTATGATTCACAAAGTAAACCTCTGTCTGT 1382
QY 233 TGAATCCACTCTTTTCCAGCTGTGTCACAAAGTCCATAGGTCATTTCTGAGC 292
DB 1381 TGGTGCATGCCGACTTCCAGCTTTAGGAAGGAATCTTGC-----TCAATTTTCA 1328
QY 293 TAACATTGAAGCCGAATAGAGGTATCGTGTGTTAACTGTTCATAGCTCCCTGCAATTG 352
DB 1327 TGGCTGGTCACTCTGCTGCTGAATCACTGGTGAGGCTTGATGAGGCTGCGAGA 1268
QY 353 GTTCCGCTGAGATGCGCACTGNTCAAAATTCGACAGATTCATCGTAGGCGCTGCCA 412
DB 1267 ACTCTGGCTTGGCTGAGCTGAGCTGCTCCACTGTGGCTGTACAGTGGGCTTTGTCCA 1208
QY 413 CGTTGTTGTTCTGTTGTTATGATCCACATAATAAATCTCTACGATCACTCTTCT 472
DB 1207 AGTAGTGGTTCTGGAATTTGATACATAATATGATCTTCTCTTTTCATCTTTGTTTC 1148
QY 473 TTCCCAACCTGGAGGTAAGGTTGTGCTCTCCCATGTGTTGAGTTCGAGCATTTATGANC 532
DB 1147 TTCCCAACCTGGTGGTAAATCCAGATGAGTAGGCAAGACAGAGAGTAGGTGTTTC 1088
QY 533 CACATAA 539
DB 1087 CTCAAAA 1081

RESULT 12
US-09-949-016-1035/c
; Sequence 1035, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
```

```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1035
; LENGTH: 2264
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-1035

Query Match      8.6%; Score 53.8; DB 4; Length 2264;
Best Local Similarity 54.9%; Pred. No. 1e-06;
Matches 128; Conservative 0; Mismatches 102; Indels 3; Gaps 1;

QY 57 GGATCTTTGAATGTTGTTCTTCTGTTATGATCAACAAGTACCTTACACCTTCACGA 116
DB 388 GGGTCCAAGTAGGTGTTCTTTATTATATGTTCAACAAAAACACTTGTCCGTTCTCA 329
QY 117 GTATATCTAATTTCCAGCCTTCTGCGAGGGTTCTTCAITCTCTAGAGCCTT---GAGTT 173
DB 328 TCAGTTTCTTGTTCCTCATCGTATGCAATCTCCTGCCACTCGTTTCTTTTCCAGTT 269
QY 174 CTTGGATCTTCCCACTGGGTTGTTTGTGTTATGATCAAAAGTAAACCTGTCTGTT 233
DB 268 TTTGGATGTTCCCACTGAGTCTTCTCTCGGTGTGATTGGCGTAGTAAACCCAGCCGTC 209
QY 234 GAATCCACTCTTTTTCAGCCTGTCGCAAGGTCATAGGTCATTTTC 286
DB 208 TTGGTGGTTCTCTCTCTCCAGCCCGGAGGAGCTGCTCTCACTGTCCTGTC 156

RESULT 13
US-08-539-205A-3/c
; Sequence 3, Application US/08539205A
; Patent No. 6001619
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Caligiuri, Maureen
; APPLICANT: Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/539,205A
; FILING DATE: 04-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2848 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 238..2535
; US-08-539-205A-3

Query Match      8.6%; Score 53.4; DB 3; Length 2848;
```


Best Local Similarity 69.9%; Pred. No. 1.6e-06; Indels 0; Gaps 0;
Matches 72; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
QY 387 ACAGATTCATCGTAGCGCGTCCACGTTGTTCTGCTGTATGATCCACATAATAA 446
Db 959 ACAGAACTTAGTTGGGGCAATCCACGTTGCTTCTGTATTATGATCTACATAATAA 900
QY 447 ACTCCTCTACGATCATCAACTCTTCTTCCCAACTGGAGGTA 489
Db 899 GTTCGGCCAAAGATTATCGGTAGCTCTTCCCATCCAGGAGGAA 857

RESULT 14
US-09-392-163A-3/c
; Sequence 3, Application US/09392163A
; Patent No. 6503742
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Caligiuri, Maureen
; APPLICANT: Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/392,163A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/539,205
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2848 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 238..2535
; US-09-392-163A-3

Query Match 8.6%; Score 53.4; DB 4; Length 2848;
Best Local Similarity 69.9%; Pred. No. 1.6e-06;
Matches 72; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
QY 387 ACAGATTCATCGTAGCGCGTCCACGTTGTTCTGCTGTATGATCCACATAATAA 446
Db 959 ACAGAACTTAGTTGGGGCAATCCACGTTGCTTCTGTATTATGATCTACATAATAA 900
QY 447 ACTCCTCTACGATCATCAACTCTTCTTCCCAACTGGAGGTA 489
Db 899 GTTCGGCCAAAGATTATCGGTAGCTCTTCCCATCCAGGAGGAA 857

RESULT 15

US-09-248-796A-4083/c
; Sequence 4083, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 4083
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Candida albicans
; US-09-248-796A-4083

Query Match 8.4%; Score 52.2; DB 4; Length 267;
Best Local Similarity 74.2%; Pred. No. 1.1e-06;
Matches 66; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
QY 402 GGCCGCTGCCACGTTGTTCTGCTGTATGATCCACATAATAAATCTCTACGATCA 461
Db 191 GGACGTTGCCAAGTTGTAGTTCTAGAAATTGTGATCAACATAATAAGTTCTACCAAAATTA 132
QY 462 TCAACTCTTCTTCCCAACTGGAGGTA 490
Db 131 TCAGTTCTACGTTCCCAACCCAGGTGGTAA 103

Search completed: June 30, 2005, 05:41:07
Job time : 162 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2005, 03:09:44 ; Search time 3088 Seconds
(without alignments)
9791.468 Million cell updates/sec

Title: US-10-618-408-3
Perfect score: 624
Sequence: 1 ccgcgcgtntaagcattt.....tccagactgctgccgnaca 624

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hgt.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	566.2	90.7	2052	6	AR162499 Sequence
c 2	566.2	90.7	2052	6	AX675520 Sequence
c 3	566.2	90.7	2052	9	HSU96113 Sequence
c 4	566.2	90.7	2827	6	AX053140 Sequence
c 5	566.2	90.7	3382	9	HSU801707 Sequence
c 6	566.2	90.7	3479	9	AY043361 Sequence
c 7	566.2	90.7	3599	9	AY345857 Sequence
c 8	566.2	90.7	3855	9	BC036065 Sequence
c 9	552.2	90.7	3855	9	BC036065 Sequence
c 10	510.8	81.9	49548	9	AC016962 Sequence
c 11	431.2	69.1	1794	5	AF412122 Sequence
c 12	367.8	58.9	3501	5	BC076842 Sequence
c 13	355.4	57.0	130034	2	AC101726 Sequence
c 14	355.4	57.0	135660	2	AC103927 Sequence
c 15	355.4	57.0	189541	2	AC101785 Sequence
c 16	355.4	57.0	233321	10	AL732546 Sequence
c 17	347	55.6	254053	2	AC109017 Sequence
c 18	255.2	40.9	1867	5	CR386634 Sequence
c 19	249.6	40.0	2565	10	AY600518 Sequence

c 20	249.2	39.9	5122	10	AF037454	AF037454 Mus muscu
c 21	249.2	39.9	5163	10	BC062934	BC062934 Mus muscu
c 22	249.2	39.9	5182	10	BC064678	BC064678 Mus muscu
c 23	246.6	39.5	2377	6	AX277586	AX277586 Sequence
c 24	246.6	39.5	2377	6	AX675519	AX675519 Sequence
c 25	246.6	39.5	2377	6	AX774812	AX774812 Sequence
c 26	246.6	39.5	2377	9	AF038564	AF038564 Homo sapi
c 27	246.6	39.5	2559	6	AR083345	AR083345 Sequence
c 28	246.6	39.5	2559	6	AR102842	AR102842 Sequence
c 29	246.6	39.5	2559	6	BD124048	BD124048 Protein.
c 30	246.6	39.5	2678	6	AR531671	AR531671 Sequence
c 31	246.6	39.5	2869	9	AB056663	AB056663 Homo sapi
c 32	246.6	39.5	2970	9	AF095745	AF095745 Homo sapi
c 33	246.6	39.5	3745	9	BC011571	BC011571 Homo sapi
c 34	246.6	39.5	5372	6	AR102841	AR102841 Sequence
c 35	246.6	39.5	5372	6	BD124047	BD124047 Protein.
c 36	246.2	39.5	5359	6	AR083344	AR083344 Sequence
c 37	233.8	37.5	180976	9	AL513321	AL513321 Human DNA
c 38	213.2	34.2	2910	10	BC039921	BC039921 Mus muscu
c 39	213.2	34.2	4319	10	BC048184	BC048184 Mus muscu
c 40	211.6	33.9	4442	9	BC013645	BC013645 Homo sapi
c 41	211.6	33.9	4500	9	BC064531	BC064531 Homo sapi
c 42	203.6	32.6	3475	6	AR162506	AR162506 Sequence
c 43	203.6	32.6	3475	6	AX675521	AX675521 Sequence
c 44	203.6	32.6	3475	9	HSU96114	HSU96114 Homo sapien
c 45	201.2	32.2	2394	6	CQ727501	CQ727501 Sequence

ALIGNMENTS

RESULT 1	AR162499/c	AR162499	Sequence 3 from patent US 6258601.	2052 bp	DNA	linear	PAT 17-OCT-2001
LOCUS	AR162499	Sequence 3 from patent US 6258601.					
DEFINITION	AR162499	Sequence 3 from patent US 6258601.					
ACCESSION	AR162499	Sequence 3 from patent US 6258601.					
VERSION	AR162499.1	GI:16229714					
KEYWORDS	Unknown.						
SOURCE	Unknown.						
ORGANISM	Unknown.						
REFERENCE	1 (bases 1 to 2052)						
AUTHORS	Monia,B.P. and Cowser,L.M.						
TITLE	Antisense modulation of ubiquitin protein ligase expression						
JOURNAL	Patent: US 6258601-A 3 10-JUL-2001;						
FEATURES	Location/Qualifiers						
source	1..2052						
	/organism="unknown"						
	/mol_type="unassigned DNA"						

ORIGIN

Query Match	90.7%	Score 566.2;	DB 6;	Length 2052;
Best Local Similarity	97.3%	Pred. No. 2.1e-144;		
Matches 582;	Conservative 0;	Mismatches 15;	Indels 1;	Gaps 1;
Qy	27	CCCTTAGTACAGTACCTCCATCCGAGGATCTTGAATGTGTGTCTGTGTTA	86	
Db	1044	CCCTTAGTACAGTACCTCCATCCGAGGATCTTGAATGTGTGTCTGTGTTA	985	
Qy	87	TGATCAACAAGTACCTTACACCTTCCAGGATATCTAATTTCCAGCCTTCTGGCAGG	146	
Db	984	TGATCAACAAGTACCTTACACCTTCCAGGATATCTAATTTCCAGCCTTCTGGCAGG	925	
Qy	147	GGTTCTTCTTGTAGGCTTGTAGGATCTTCCACCTGGGTGTGTGTGTGTTA	206	
Db	924	GGTTCTTCTTGTAGGCTTGTAGGATCTTCCACCTGGGTGTGTGTGTGTTA	865	
Qy	207	TGATTCACAAAGTAAACCTGTCTGTGTAATCACTCTTTTCCAGCTGTGTGCAAA	266	
Db	864	TGATTCACAAAGTAAACCTGTCTGTGTAATCACTCTTTTCCAGCTGTGTGCAAA	805	
Qy	267	GGTCCATAGGGTTCATTTCTGAGCTACATTTGAAGCCGATAGAGGTATCGTTGTTA	326	

804 GGTCCATTAAGGTCATTTCTGCGCTAACAATTGAAGCCGAATAGAGGTATCGTTGGTTA 745
 327 AACTGTTGATAGTCCCTCGCAATTGGTTCGCTGAGATTGCACTGNTCAAAATTCGG 386
 744 AACTGTTGATAGTCCCTCGCAATTGGTTCGCTGAGATTGCACTGNTCAAAATTCGG 685
 387 ACAGATTCCATGTTAGGCGCTGCCACGTTGTTGTTCTGGTCTTAAGTACCAATAATAA 446
 684 ACAGATTCCATGTTAGGCGCTGCCACGTTGTTGTTCTGGTCTTAAGTACCAATAATAA 625
 447 ACTCCTCTACGATCATCAACTCTTCTTCCCACTGGAGGTAAGGTTGTGCTCTCTCC 506
 624 ACTCCTCTACGATCATCAACTCTTCTTCCCACTGGAGGTAAGGTTGTGCTCTCTCC 565
 507 CATGTGTTAGTTCGAGCATTTATGATCCACATATAGCTCTACCATGAGGATCTTTTCTT 566
 564 CATGTGTTAGTTCGAGCATTTATGATCCACATATAGCTCTACCATGAGGATCTTTTCTT 505
 567 TGTTCCTCCACCTGATGGCAANGTTCTGTGTTGGCAATTCCTCCAGACTGCTGCGGNACA 624
 504 TGTTCCTCA-CCCTGATGGCAAGGTTCTGTGTTGGCAATTCCTCCAGACTGCTGCGGTACA 448

RESULT 2
 AX675520/c
 LOCUS AX675520 2052 bp DNA linear PAT 27-MAR-2003
 DEFINITION Sequence 3 from Patent WO02090549.
 ACCESSION AX675520
 VERSION AX675520.1 GI:29333538
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 Greener, T., Moskowitz, H., Reiss, Y. and Alroy, I.
 REFERENCES Compositions and methods for the modulation of viral maturation
 Patents: WO 02090549-A 3 14-NOV-2002;
 Proteologics, Ltd. (US)
 TITLE Location/Qualifiers
 JOURNAL 1..2052
 FEATURES
 source
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Query Match 90.7%; Score 566.2; DB 6; Length 2052;
 Best Local Similarity 97.3%; Pred. No. 2.1e-144;
 Matches 592; Conservative 0; Mismatches 15; Indels 1; Gaps 1;
 27 CCCTTTAGTNCAGATGACTTCCCATGCGAGGATCTTGAATGTTGTTGTTCTTGTGTTA 86
 1044 CCTTTAGTTACAGATGACTTCCCATGCGAGGATCTTGAATGTTGTTGTTCTTGTGTTA 985
 87 TGAATCAACAAGTACCTTACACTTCCAGAGTATCTAAATTCCTCCAGCTTCTGCGAGG 146
 984 TGAATCAACAAGTACCTTACACTTCCAGAGTATCTAAATTCCTCCAGCTTCTGCGAGG 925
 147 GGTTCCTCAATCTGTAAGCTTGAAGTCTTGAATCTTCCACTGCGAGGTTGTTGTTGTTA 206
 924 GGTTCCTCAATCTGTAAGCTTGAAGTCTTGAATCTTCCACTGCGAGGTTGTTGTTGTTA 865
 207 TGAATCAACAAGTAAACCTTCTGTTGAATCACTCTTTTTCCTCCAGCTGCTGCGAGG 266
 864 TGAATCAACAAGTAAACCTTCTGTTGAATCACTCTTTTTCCTCCAGCTGCTGCGAGG 805
 267 GGTCCATAAGGTCATTTTCTGAGCTAACTTGAAGCGAATAGAGGTATCGTTGGTTA 326
 804 GGTCCATAAGGTCATTTTCTGAGCTAACTTGAAGCGAATAGAGGTATCGTTGGTTA 745
 327 AACTGTTGATAGTCCCTCGCAATTGGTTCGCTGAGATTGCACTGNTCAAAATTCGG 386
 744 AACTGTTGATAGTCCCTCGCAATTGGTTCGCTGAGATTGCACTGNTCAAAATTCGG 685

387 ACAGATTCCATGTTAGGCGCTGCCACGTTGTTGTTCTGGTCTTAAGTACCAATAATAA 446
 684 ACAGATTCCATGTTAGGCGCTGCCACGTTGTTGTTCTGGTCTTAAGTACCAATAATAA 625
 447 ACTCCTCTACGATCATCAACTCTTCTTCCCACTGGAGGTAAGGTTGTGCTCTCTCC 506
 624 ACTCCTCTACGATCATCAACTCTTCTTCCCACTGGAGGTAAGGTTGTGCTCTCTCC 565
 507 CATGTGTTAGTTCGAGCATTTATGATCCACATATAGCTCTACCATGAGGATCTTTTCTT 566
 564 CATGTGTTAGTTCGAGCATTTATGATCCACATATAGCTCTACCATGAGGATCTTTTCTT 505
 567 TGTTCCTCCACCTGATGGCAANGTTCTGTGTTGGCAATTCCTCCAGACTGCTGCGGNACA 624
 504 TGTTCCTCA-CCCTGATGGCAAGGTTCTGTGTTGGCAATTCCTCCAGACTGCTGCGGTACA 448

RESULT 3
 HSU96113/c
 LOCUS HSU96113 2052 bp mRNA linear PRI 29-MAY-1997
 DEFINITION Homo sapiens Nedd-4-like ubiquitin-protein ligase WWP1 mRNA, partial cds.
 ACCESSION U96113
 VERSION U96113.1 GI:2072500
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 Pirozzi, G., McConnell, S.J., Uveges, A.J., Carter, J.M., Sparks, A.B.,
 Kay, B.K. and Fowlkes, D.M.
 REFERENCES Identification of novel human WW domain-containing proteins by
 cloning of ligand targets
 J. Biol. Chem. 272 (23), 14611-14616 (1997)
 97313427
 MEDLINE 2 (bases 1 to 2052)
 PUBMED 9169421
 Pirozzi, G. and Uveges, A.
 TITLE Direct Submission
 JOURNAL Submitted (02-APR-1997) CytoGen Corp., 201 College Road East,
 Princeton, NJ 08540, USA
 FEATURES
 source
 1..2052
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 <1..>2052
 /notes="Nedd-4-like ubiquitin-protein ligase; WW
 domain-containing protein"
 /codon_start=2
 /product="WWP1"
 /protein_id="AAC51324.1"
 /db_xref="GI:2072501"
 /translation="TNHPTSTLVQNSCCSYVNGDNTSPSSQVAARPKNTPAPKPL
 ASPADLVNGSSSPAPDASVTGTPVSEENALSPNCTSTTVDPVQQLITSE
 NNECISTSELESEARSILEPDTNSRSSSAPEAKSRQPDGMDPVQSGNANTE
 TLPSEQRNDPHGRITTYVDHNTTWTWERPOPLPGWERRVDRRRVYVDHNTTT
 TWQPTMESVRQDSQORNLQGANQFNRKYLXSASMLAENDPPGLPGMEKR
 VSDTRVYFNHNTKTQWEDPRQGLQNEEPLPEGWETRYREGVRYFVDHNTT
 FKDPNGKSVTKGPGQIAYERGRWKLHFRYLQCSNALPSHVKLINVSQTLPEDSF
 QQIMALKPYDLRRLLYVIFRGEGLDYGGLAREWFFLLSHEVLNPMYCLFEYAGKNY
 CLQINPASTINPDLVSFCFPIGRFAMALFHGKFIIDTGFSLDKLGSNLIIVTEENDE
 SITDEFNSLIIVTDNNIEKGLMEYFSDMEILGKVTIDTGFSLDKLGSNLIIVTEENDE
 YIGLMTWRFSGVQGTAKFLDGFNEVPLQWLOVDFEKELEMLCGMQEVDLADWQ
 RNTYRHYTRNSKQILNFWQFVKETDNEVRMLLQFVTGTGTCRLPLGFAELMGSGNPR
 NS"

ORIGIN
 Query Match 90.7%; Score 566.2; DB 9; Length 2052;
 Best Local Similarity 97.3%; Pred. No. 2.1e-144;
 Matches 582; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

```
QY 27 CCCTTTAGTACAGATGACTTCCCATTCGAGGATCTTTGAATGTTGTTGTTCTTTGTGTTA 86
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1044 CCTTTAGTTACAGATGACTTCCCATTCGAGGATCTTTGAATGTTGTTGTTCTTTGTGTTA 985
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 87 TGATCAACAAGTACCTTTACACCTTCACGAGTATATCTAAATTTCCACAGCCTTCTGGCAGG 146
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 984 TGATCAACAAGTACCTTTACACCTTCACGAGTATATCTAAATTTCCACAGCCTTCTGGCAGG 925
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 147 GGTTCCTTCATTCTGTAAAGCCCTTCAGTCTCTTGGATCTTCCCACTGGGTTGTTTTGTGTTA 206
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 924 GGTTCCTTCATTCTGTAAAGCCCTTCAGTCTCTTGGATCTTCCCACTGGGTTGTTTTGTGTTA 865
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 207 TGATTCACAAAGTAAACCCCTGCTGTGTTGAATCCACTCTTTTTCCACAGCCTGTTGGCAAA 266
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 864 TGATTCACAAAGTAAACCCCTGCTGTGTTGAATCCACTCTTTTTCCACAGCCTGTTGGCAAA 805
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 267 GGTCCCAATAGGGTCATTTCTGCGAGCTAAACATTTGAAGCCCAATAGAGGTATCGTTGGTTA 326
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 804 GGTCCCAATAGGGTCATTTCTGCGAGCTAAACATTTGAAGCCCAATAGAGGTATCGTTGGTTA 745
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 327 AACTGTTGTCATAGCTCCCTGCAATTTGGTTCGGCTGAGATTCACCACTGNTCAAAATTCGG 386
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 744 AACTGTTGTCATAGCTCCCTGCAATTTGGTTCGGCTGAGATTCACCACTGNTCAAAATTCGG 685
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 387 ACAGATTCATAGTGGCCGCTGCGACGTTGTTCTGTTGTTCTGTTGTTATGATCCACATAATA 446
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 684 ACAGATTCATAGTGGCCGCTGCGACGTTGTTCTGTTGTTCTGTTGTTATGATCCACATAATA 625
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 447 ACTCCTCTACGATCATCACTCTTTTCCCACTGGAGTAAAGGTTGTTGTTCTCTCC 506
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 624 ACTCCTCTACGATCATCACTCTTTTCCCACTGGAGTAAAGGTTGTTGTTCTCTCC 565
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 507 CATGTGTTAGTTCGAGATTTATGATCCACATAATAGCTCTACCATGAGGATCTTTTCTT 566
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 564 CATGTGTTAGTTCGAGATTTATGATCCACATAATAGCTCTACCATGAGGATCTTTTCTT 505
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 567 TGTTCACACCCCTGATGGCAAGNTTCTGTGTTGGCATTTCCAGACTGCTGCCGNACA 624
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 504 TGTTCACCA-CCCTGATGGCAAGGTTCTGTGTTGGCATTTCCAGACTGCTGCCGNACA 448
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 4
AX053140/c
LOCUS AX053140 2827 bp DNA linear PAT 13-JAN-2001
DEFINITION Sequence 2 from Patent WO0073329.
ACCESSION AX053140
VERSION AX053140.1 GI:12227505
KEYWORDS
SOURCE Drosophila sp.
ORGANISM Drosophila sp.
REFERENCE Baron, M.
AUTHORS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
TITLE Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
JOURNAL Ephydroidea; Drosophilidae; Drosophila.
FEATURES
source
1. .2827
/organism="Drosophila sp."
/mol_type="unassigned DNA"
/db_xref="taxon:7242"

ORIGIN
Query Match 90.7%; Score 566.2; DB 6; Length 2827;
Best Local Similarity 97.3%; Pred. No. 2.1e-144;
Matches 582; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

QY 27 CCCTTTAGTACAGATGACTTCCCATTCGAGGATCTTTGAATGTTGTTGTTCTTTGTGTTA 86
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1043 CCTTTAGTTACAGATGACTTCCCATTCGAGGATCTTTGAATGTTGTTGTTCTTTGTGTTA 984
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
QY 87 TGATCAACAAGTACCTTTACACCTTCACGAGTATATCTAAATTTCCACAGCCTTCTGGCAGG 146
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 983 TGATCAACAAGTACCTTTACACCTTCACGAGTATATCTAAATTTCCACAGCCTTCTGGCAGG 924
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 147 GGTTCCTTCATTCTGTAAAGCCCTTCAGTCTCTTGGATCTTCCCACTGGGTTGTTTTGTGTTA 206
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 923 GGTTCCTTCATTCTGTAAAGCCCTTCAGTCTCTTGGATCTTCCCACTGGGTTGTTTTGTGTTA 864
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 207 TGATTCACAAAGTAAACCCCTGCTGTGTTGAATCCACTCTTTTTCCACAGCCTGTTGGCAAA 266
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 863 TGATTCACAAAGTAAACCCCTGCTGTGTTGAATCCACTCTTTTTCCACAGCCTGTTGGCAAA 804
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 267 GGTCCCAATAGGGTCATTTCTGCGAGCTAAACATTTGAAGCCCAATAGAGGTATCGTTGGTTA 326
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 803 GGTCCCAATAGGGTCATTTCTGCGAGCTAAACATTTGAAGCCCAATAGAGGTATCGTTGGTTA 744
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 327 AACTGTTGTCATAGCTCCCTGCAATTTGGTTCGGCTGAGATTCACCACTGNTCAAAATTCGG 386
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 743 AACTGTTGTCATAGCTCCCTGCAATTTGGTTCGGCTGAGATTCACCACTGNTCAAAATTCGG 684
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 387 ACAGATTCATAGTGGCCGCTGCGACGTTGTTCTGTTGTTCTGTTGTTATGATCCACATAATA 446
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 683 ACAGATTCATAGTGGCCGCTGCGACGTTGTTCTGTTGTTCTGTTGTTATGATCCACATAATA 624
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 447 ACTCCTCTACGATCATCACTCTTTTCCCACTGGAGTAAAGGTTGTTGTTCTCTCC 506
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 623 ACTCCTCTACGATCATCACTCTTTTCCCACTGGAGTAAAGGTTGTTGTTCTCTCC 564
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 507 CATGTGTTAGTTCGAGATTTATGATCCACATAATAGCTCTACCATGAGGATCTTTTCTT 566
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 563 CATGTGTTAGTTCGAGATTTATGATCCACATAATAGCTCTACCATGAGGATCTTTTCTT 504
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 567 TGTTCACACCCCTGATGGCAAGNTTCTGTGTTGGCATTTCCAGACTGCTGCCGNACA 624
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 503 TGTTCACCA-CCCTGATGGCAAGGTTCTGTGTTGGCATTTCCAGACTGCTGCCGNACA 447
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
RESULT 5
HSM801707/c
LOCUS HSM801707 3382 bp mRNA linear PRI 20-MAR-2002
DEFINITION Homo sapiens mRNA; cDNA DKFZp434D2111 (from clone DKFZp434D2111).
ACCESSION AL136739
VERSION AL136739.1 GI:12052996
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 3382)
AUTHORS Duesterhoeft, A., Lauber, J., Mewes, H.W., Weil, B. and Wiemann, S.
TITLE Direct Submission
JOURNAL Submitted (12-MAR-2002) MIPS, Am Klopferspitz 18a, D-82152
Martinsried, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
consortium of the German Genome Project.
This clone (DKFZp434D2111) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.
FEATURES
Location/Qualifiers
1. .3382
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp434D2111"
/db_xref="taxon:9606"
/chromosome="3"
/clone="DKFZp434D2111"
/tissue_type="testis"
/clone_lib="434 (synonym: htes3). Vector pSport1; host
```

```

DH10B; sites NotI + SalI"
/dev_stage="adult"
1..3382
/gene="DKFZp434D2111"
1..2779
/gene="DKFZp434D2111"
/notes="similarity to Nedd-4-like ubiquitin-protein ligase
(Homo sapiens)"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB66673.1"
/db_xref="GI:12052997"
/db_xref="GOA:Q9H0M0"
/db_xref="UniProt/TrEMBL:Q9H0M0"
/translation="NATASPRSDTSNNHSGRLQLQVTVSSAKLRKKQWFGTAIYVE
VDGEGITTKAKSSSNPKWDQLTVNVPQTTLFQWMSHRLKADALLGKATIDLK
QALLHNRLKRVKEQLKLENGKIAQTLVLDGLVIEQENITNCSSPTIEI
QENGDAHNGECPARTTARLAVEGTNGIDNHVPTSTLVONSCCSVVNGDNTSPSP
QVAAPKNTAPKPLASEPADDTVNGESSFAPTDNASVTGTPVVEENALSPNCTST
TVEDPVPQILTSSENNECIPSTSAELEARLLEPDTNSRSSSAPEAAKSRQPDG
CNDPVRQSGNANTETLPQSGWQRKDPHGRTYYVDHNTKTTTWERPQPLPGWERRVD
DRRVYVVDHNTTTLTQWPTMESVRNFQWQSRNQLQGMQFNQRLYSASMLAA
ENDPYGLPGEWKEKRVSDTRVYFVNHNKTQWEDPRTQGLQNEEPLPEGWEIYR
EGVRYFVDHNTTTLTQWPTMESVRNFQWQSRNQLQGMQFNQRLYSASMLAA
VKINVSQTLFEDSFQIIMALKPYDLRLRYLVIERGEGLDYGLAREWFFLLSHVL
NPMYCLFEYAGKNNYCLQINPASTINPDHLSYFCFGRFIAMALHKGFDITGFSLP
YKMLSKLTIKDLSEIDTFYNSLIWIHNNIECGLEMYFSVMEILLKVTSHDLK
LGGSNLTVEENKDEYIGLMTWFRSGVQBOETKAFDGFNEVPLQWQLQYDEKLE
VLMCGMQEVDLADWQNTVYRHTYRNSKQIIFWQFVKETDNEVRMLQFVTGRL
PLGGFAELMGSNGPQKFCIEKVGKDTLWLRSHSTCFNRLDLPPYKSYEQLEKLLFAIE
ETSGFQGE"
polyA_signal 3346..3351
/gene="DKFZp434D2111"
polyA_site 3362
/gene="DKFZp434D2111"
ORIGIN
Query Match 90.78; Score 566.2; DB 9; Length 3382;
Best Local Similarity 97.34; Pred. No. 2e-144; Indels 1; Gaps 1;
Matches 582; Conservative 0; Mismatch 15;
QY 27 CCCTTTAGTCAGATGACTTCCCATTCGCGAGATCTTTGAATGTTGTTGTTGTTA 86
DB 1620 CCTTTAGTCAGATGACTTCCCATTCGCGAGATCTTTGAATGTTGTTGTTGTTA 1561
QY 87 TGATCAACAAAGTACCTTACCTTCAGGATATATCTAATTTCCAGCCTTCGCGAG 146
DB 1560 TGATCAACAAAGTACCTTACCTTCAGGATATATCTAATTTCCAGCCTTCGCGAG 1501
QY 147 GGTCTTCATCTCTAGCCCTTGAGTCTTGGATCTTCCACTGGGTGTTTGTGTTA 206
DB 1500 GGTCTTCATCTCTAGCCCTTGAGTCTTGGATCTTCCACTGGGTGTTTGTGTTA 1441
QY 207 TGATTCACAAAGTAAACCTGCTCTGTTGAATCCACTCTTTTTTCCAGCCTGTCGCAA 266
DB 1440 TGATTCACAAAGTAAACCTGCTGTTGAATCCACTCTTTTTTCCAGCCTGTCGCAA 1381
QY 267 GGTCAATAGGTCATTTTTCGACGCTAACATTCGACCCGAATAGAGTATCGTTGTTA 326
DB 1380 GGTCAATAGGTCATTTTTCGACGCTAACATTCGACCCGAATAGAGTATCGTTGTTA 1321
QY 327 AACTGTTCCATAGCTCCCTGCAATTTGGTTCGCTGAGATTCGCCACTGNATCAAAATTNCGG 386
DB 1320 AACTGTTCCATAGCTCCCTGCAATTTGGTTCGCTGAGATTCGCCACTGNATCAAAATTNCGG 1261
QY 387 ACAGATTCATGTTAGGCGCGTGCACATGTTGTTGTTGTTGTTATGATCACAATAATA 446
DB 1260 ACAGATTCATGTTAGGCGCGTGCACATGTTGTTGTTGTTGTTATGATCACAATAATA 1201
QY 447 ACTCCTCTACGATCATCACTCTTCTTCCCACTCGAGGTAAAGGTTGTTGTTCTCTCC 506
DB 1200 ACTCCTCTACGATCATCACTCTTCTTCCCACTCGAGGTAAAGGTTGTTGTTCTCTCC 1141
```

```

QY 507 CATGTGGTAGTTCGAGCATTATGANCACATAATACGCTCTTACCATGAGATCTTTCTT 566
DB 1140 CATGTGGTAGTTCGAGTATTATGATCCACATAATAGGTTCTTACCATGAGATCTTTCTT 1081
QY 567 TGTTCACACCCCTGATGCGAANGNTTCTGTGTGGCATTCACAGACTGCTGCCGNACA 624
DB 1080 TGTTCACA-CCCTGATGCGAAGGTTCTGTGTGGCATTCACAGACTGCTGCCGNACA 1024

RESULT 6
AY043361/c 3479 bp mRNA linear PRI 01-SEP-2001
LOCUS Homo sapiens WW domain-containing protein 1 mRNA, complete cds.
DEFINITION AY043361
ACCESSION AY043361.1 GI:15419010
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3479)
AUTHORS Flasa,M., Aslam,H., Poulson,R., Hanby,A., Gorman,P., Roylance,R.,
Nolan,T., Brady,G., Canfield,A. and Baron,M.
WP1, a human homolog of Drosophila suppressor of deltex
TITLE Unpublished
JOURNAL
REFERENCE 2 (bases 1 to 3479)
AUTHORS Flasa,M., Aslam,H., Poulson,R., Hanby,A., Gorman,P., Roylance,R.,
Nolan,T., Brady,G., Canfield,A. and Baron,M.
Direct Submission
TITLE Submitted (03-JUL-2001) Biological Sciences, University of
JOURNAL Manchester, Stoford Building, Oxford Rd., Manchester M13 9PT, UK
FEATURES
source
1..3479
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosomes="8"
/map="8q21"
86..2854
/notes="WWP1; Nedd4 family E3 ubiquitin ligase; suppressor
of deltex related protein 1 (hSDRP1)"
/codon_start=1
/product="WW domain-containing protein 1"
/protein_id="AAK94668.1"
/db_xref="GI:15419011"
/translation="NATASPRSDTSNNHSGRLQLQVTVSSAKLRKKQWFGTAIYVE
VDGEGITTKAKSSSNPKWDQLTVNVPQTTLFQWMSHRLKADALLGKATIDLK
QALLHNRLKRVKEQLKLENGKIAQTLVLDGLVIEQENITNCSSPTIEI
QENGDAHNGECPARTTARLAVEGTNGIDNHVPTSTLVONSCCSVVNGDNTSPSP
QVAAPKNTAPKPLASEPADDTVNGESSFAPTDNASVTGTPVVEENALSPNCTST
TVEDPVPQILTSSENNECIPSTSAELEARLLEPDTNSRSSSAPEAAKSRQPDG
CNDPVRQSGNANTETLPQSGWQRKDPHGRTYYVDHNTKTTTWERPQPLPGWERRVD
DRRVYVVDHNTTTLTQWPTMESVRNFQWQSRNQLQGMQFNQRLYSASMLAA
ENDPYGLPGEWKEKRVSDTRVYFVNHNKTQWEDPRTQGLQNEEPLPEGWEIYR
EGVRYFVDHNTTTLTQWPTMESVRNFQWQSRNQLQGMQFNQRLYSASMLAA
VKINVSQTLFEDSFQIIMALKPYDLRLRYLVIERGEGLDYGLAREWFFLLSHVL
NPMYCLFEYAGKNNYCLQINPASTINPDHLSYFCFGRFIAMALHKGFDITGFSLP
YKMLSKLTIKDLSEIDTFYNSLIWIHNNIECGLEMYFSVMEILLKVTSHDLK
LGGSNLTVEENKDEYIGLMTWFRSGVQBOETKAFDGFNEVPLQWQLQYDEKLE
VLMCGMQEVDLADWQNTVYRHTYRNSKQIIFWQFVKETDNEVRMLQFVTGRL
PLGGFAELMGSNGPQKFCIEKVGKDTLWLRSHSTCFNRLDLPPYKSYEQLEKLLFAIE
ETSGFQGE"
misc_feature 137..508
/notes="Region: C2 domain"
misc_feature 1130..1222
/notes="Region: WW domain"
misc_feature 1223..1333
/notes="Region: WW domain"
misc_feature 1451..1543
/notes="Region: WW domain"
misc_feature 1571..1666
/notes="Region: WW domain"
misc_feature 1787..2851
/notes="Region: WW domain"
```

ORIGIN	/note="Region: HECT domain"		source		
Query Match	90.7%; Score 566.2; DB 9; Length 3479;				
Best Local Similarity	97.3%; Pred. No. 2e-144;				
Matches 582; Conservative	0; Mismatches 15; Indels 1; Gaps 1;				
QY	27	CCCTTTAGTACAGATGCTCCCAATCGAGGATCTTTGAATGTTGTTGTTGTTA 86			
Db	1695	CCTTTAGTACAGATGCTCCCAATCGAGGATCTTTGAATGTTGTTGTTGTTA 1636			
QY	87	TGATCAACAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGCCTTCGGCAGG 146			
Db	1635	TGATCAACAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGCCTTCGGCAGG 1576			
QY	147	GGTTCCTTCATCTGTAAAGCCTTCAGTCTTGGATCTTCCCACTGGGTGTTTGTGTTA 206			
Db	1575	GGTTCCTTCATCTGTAAAGCCTTCAGTCTTGGATCTTCCCACTGGGTGTTTGTGTTA 1516			
QY	207	TGATTCACAAGTAAACCTGTCTGTTGAATFCACTCTTTTTCAGCCTTCGGTGGCAA 266			
Db	1515	TGATTCACAAGTAAACCTGTCTGTTGAATFCACTCTTTTTCAGCCTTCGGTGGCAA 1456			
QY	267	GGTCCATAAGGTCATTTCTGCAGCTAACTTGAAGCCGAATAGAGGTATCGTTGTTA 326			
Db	1455	GGTCCATAAGGTCATTTCTGCAGCTAACTTGAAGCCGAATAGAGGTATCGTTGTTA 1396			
QY	327	AACGTGTGATAGCTCCCTGCAATGGTTCGGTGCAGATTCGCACTGNTCAAAATTCGG 386			
Db	1395	AACGTGTGATAGCTCCCTGCAATGGTTCGGTGCAGATTCGCACTGNTCAAAATTCGG 1336			
QY	387	ACAGATTCATAGTGGAGCCTGCGACGTTGTTGTTGTTGTTGTTGTTGTTA 446			
Db	1335	ACAGATTCATAGTGGAGCCTGCGACGTTGTTGTTGTTGTTGTTGTTGTTA 1276			
QY	447	ACTCTCTACGATCATCAACTCTTCTTCCCACTGGAGGTAAAGGTGTTGTTCTCTCC 506			
Db	1275	ACTCTCTACGATCATCAACTCTTCTTCCCACTGGAGGTAAAGGTGTTGTTCTCTCC 1216			
QY	507	CATGTGTAGTTCGAGCATTTATGACCACATATACCTCTACCATGAGGATCTTTTCTT 566			
Db	1215	CATGTGTAGTTCGAGCATTTATGATCCACATATAGTTCTACCATGAGGATCTTTTCTT 1156			
QY	567	TGTTCCCACTGATGGCAANGNTCTGTGTTGGCATTCCTCCAGACTGCTGCCGNACA 624			
Db	1155	TGTTCCCA-CCCTGATGGCAAGGTTCTGTGTTGGCATTCCTCCAGACTGCTGCCGTACA 1099			
RESULT 7					
AY345857/c					
LOCUS					
DEFINITION	Homo sapiens TGIF-interacting ubiquitin ligase 1 mRNA, complete cds.				
ACCESSION	AY345857				
VERSION	AY345857.1				
KEYWORDS	GI:33590475				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
	1 (bases 1 to 3599)				
REFERENCE	Seo, S.R., Lallemand, F., Ferrand, N., Pessah, M., L'Hoste, S.,				
AUTHORS	Camonis, J. and Attfi, A.				
TITLE	The novel E3 ubiquitin ligase Tiul1 associates with TGIF to target				
	Snad2 for ubiquitin-mediated degradation				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 3599)				
AUTHORS	Seo, S.R., Lallemand, F., Ferrand, N., Pessah, M., L'Hoste, S.,				
	Camonis, J. and Attfi, A.				
TITLE	Direct Submission				
JOURNAL	Submitted (18-JUL-2003) INSERM U482, Hopital Saint-Antoine, 184 Rue				
	du Faubourg St-Antoine, Paris 75571, France				
	Location/Qualifiers				
FEATURES					

RESULT 8	BC036065/c	3855 bp	mrna	linear	PRI 19-JUL-2004
LOCUS	BC036065				
DEFINITION	Homo sapiens WW domain containing E3 ubiquitin protein ligase 1, mRNA (cDNA clone MGC:33686 IMAGE:5296005), complete cds.				
ACCESSION	BC036065				
VERSION	1				
KEYWORDS	BC036065.1	GI:23271280			
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. (bases 1 to 3855)				
AUTHORS	Strausberg, R.L., Feingold, B.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D., Aklonis, S.F., Zeeberg, B., Moore, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshikiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, D.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalios, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.				
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences				
PUBLISHED	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)				
PUBMED	12477932				
REFERENCE	2 (bases 1 to 3855)				
AUTHORS	Strausberg, R.				
TITLE	Direct Submission				
JOURNAL	Submitted (31-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov				
COMMENT	Contact: MGC help desk Email: cgaps-r@mail.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN) cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: http://www-ahgc.stanford.edu Contact: (Dickson, Mark) mdc@paxil.stanford.edu Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.				
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IPAK Plate: 48 Row: K Column: 20 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 19718739. Location/Qualifiers 1..3855 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:33686 IMAGE:5296005" /tissue_type="Testis" /clone_lib="NIH_MGC_97" /lab_host="DH10B" /note="Vector: pBluescript" 1..3855				
source					
gene					

gene

Db 1354 TGTTCCTCA-CCTGATGCGAAGGTTCTGTGTGGCATTCCTCCAGACTGCTGCCGTACA 1298
|||||
RESULT 9
AX053142/c
LOCUS
DEFINITION Sequence 4 from Patent WO0073329.
ACCESSION AX053142
VERSION AX053142.1 GI:12227506
KEYWORDS
SOURCE Drosophila sp.
ORGANISM Drosophila sp.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscophora;
Ephydroidea; Drosophilidae; Drosophila.
1
REFERENCE
AUTHORS Baron, M.
TITLE Uses of notch related genes
JOURNAL Patent: WO 0073329-A 4 07-DEC-2000;
THE VICTORIA UNIVERSITY OF MANCHESTER (GB)
FEATURES
source Location/Qualifiers
1..3494
/organism="Drosophila sp."
/mol_type="unassigned DNA"
/db_xref="taxon:7242"
ORIGIN
Query Match 88.58; Score 552.2; DB 6; Length 3494;
Best Local Similarity 97.08; Pred. No. 1.4e-140;
Matches 580; Conservative 0; Mismatches 15; Indels 3; Gaps 2;
QY 27 CCCTTTAGTNCAGATGACTTCCCATTCGAGGATCTTTGAATGTTGTTCTTGTGTTA 86
|||
Db 1710 CCTTTAGTACAGATGACTTCCCATTCGAGGATCTTTGAATGTTGTTCTTGTGTTA 1653
|||
QY 87 TGATCAACAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGCCTTCCTGGCAGG 146
|||
Db 1652 TGATCAACAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGCCTTCCTGGCAGG 1593
|||
QY 147 GGTTCCTCATCTGTAGACCTTGGATCTTGGATCTTCCACCTGGTGTGTTTGTGTTA 206
|||
Db 1592 GGTTCCTCATCTGTAGACCTTGGATCTTGGATCTTCCACCTGGTGTGTTTGTGTTA 1533
|||
QY 207 TGATTCACAAAGTAAACCTCTGTGTTGAATCCACTCTTTTCCAGCCTGGTGGCAAA 266
|||
Db 1532 TGATTCACAAAGTAAACCTCTGTGTTGAATCCACTCTTTTCCAGCCTGGTGGCAAA 1473
|||
QY 267 GGTCCATAAGGGTCATTTCTGCAGCTTAAACATTGAAGCCGAATAGAGGTATCGTTGGTTA 326
|||
Db 1472 GGTCCATAAGGGTCATTTCTGCAGCTTAAACATTGAAGCCGAATAGAGGTATCGTTGGTTA 1413
|||
QY 327 AACTGTTGTCATGCTCCCTGCATTTGGTTCGGCTGAGATTCGCCACTGNTCAAAATTCGG 386
|||
Db 1412 AACTGTTGTCATGCTCCCTGCATTTGGTTCGGCTGAGATTCGCCACTGTTCAAAATTCGG 1353
|||
QY 387 ACAGATTCATGTTAGCGCTGCGACGTTGTTCTGTTGTTATGATCCACATATAA 446
|||
Db 1352 ACAGATTCATGTTAGCGCTGCGACGTTGTTCTGTTGTTATGATCCACATATAA 1293
|||
QY 447 ACTCCTCTACATCATCAACTCTTTTCCCAACCTCGAGGTAAAGTTGTGTCCTCC 506
|||
Db 1292 ACTCCTCTACATCATCAACTCTTTTCCCAACCTCGAGGTAAAGTTGTGTCCTCC 1233
|||
QY 507 CATGTGTGATTCGAGCATTTATGANCACATATACCTCTACCATGAGGATCTTTCTT 566
|||
Db 1232 CATGTGTGATTCGAGCATTTATGATCCACATATAGTTCTTACCATGAGGATCTTTCTT 1173
|||
QY 567 TGTTCCTCACCTCATGTCGAAGGTTCTGTGTTGGCATTCCTCCAGCTGCTGCCGAC 624
|||
Db 1172 TGTTCCTCA-CCTGATGCGAAGGTTCTGTGTGGCATTCCTCCAGACTGCTGCCGTACA 1116
|||||

RESULT 10
AC016962
LOCUS
DEFINITION Homo sapiens 3 BAC RP11-569H14 (Roswell Park Cancer Institute Human BAC Library) complete sequence.
ACCESSION AC016962.38 GI:21206149
VERSION
KEYWORDS
SOURCE HTG.
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 49548)
REFERENCE
AUTHORS Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C., Alsbrooks, S.L., Amaratunge, H.C., Are, J.R., Ayele, M., Banks, T., Barbara, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyie, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harries, C., Harris, K., Hart, M., Havlak, P., Hawes, A., He, X., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Honsi, F., Howard, S., Huber, J., Hui, K.S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loubege, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Moore, S., Morgan, M., Moorish, T., Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Nguyen, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okunolu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoshitari, N., Sisson, I., Sodergren, E., Sonaite, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Taylor, T., Tellford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wlezyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Naylor, S.L., Weinstein, G. and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 49548)
Worley, K.C.
Direct Submission
Submitted (09-DEC-1999) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 49548)
Worley, K.C.
Direct Submission
Submitted (09-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 49548)
Worley, K.C.
Direct Submission
Submitted (25-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA
5 (bases 1 to 49548)
Worley,K.C.
Direct Submission
Submitted (29-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
6 (bases 1 to 49548)
Worley,K.C.
Direct Submission
Submitted (30-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 25, 2002 this sequence version replaced gi:20279350.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
Features listing.

ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550) searches
of a local database that includes entries from dbSTS, GDB, and
local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green,
unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
EST and cDNA sequences. Genes demonstrate at least two exons
flanked by consensus splice sites that maintained sequence
continuity across the splice junctions. Sequences that are not
identical matches are annotated as similar.

SEQUENCING READ COVERAGE:Sequencing is completed to a minimum
standard of double strand coverage with a minimum of 2 clones and 2
reads with no ambiguities or 2 chemistries with a minimum of 2
clones and 3 reads with no ambiguities. If the sequence quality for
a region does not meet this standard, it will be indicated in the
annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality
standards - estimated error rate less than 1 per 10,000 bases.
Reports of lowest quality individual bases and measures of base
quality are listed below. Description of the metrics can be found
at URL:
<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT.

FEATURES	Location/Qualifiers
source	1..49548 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="3" /clone="RP11-569H14"
repeat_region	1..311 /rpt_family="AluSp"
repeat_region	506..808 /rpt_family="AluSg"
repeat_region	1125..1421 /rpt_family="AluSg"
repeat_region	complement(11498..1798) /rpt_family="AluSg"
repeat_region	complement(2205..2508) /rpt_family="AluSg"
repeat_region	complement(2606..3084) /rpt_family="MLTID"
repeat_region	3748..3897 /rpt_family="AluJb"

QY 207 TGATTCAAAAGTAAACCCCTGCTGTGGTGAATCCACTCTTTTCCCGAGCCCTGGTGGCAAA 266
 Db 45182 CGATTCAAAAGTAAACCCCTGCTGTGGTGAATCCACTCTTTTCCCGAGCCCTGGTGGCAAA 45241
 QY 267 GGTCCATAAGGCTCATTTTCTGACGCTAAACATTGAAGCGAATAGAGGTATCGTTGGTTA 326
 Db 45242 GGTCCATAAGGCTCATTTTCTGACGCTAAACATTGAAGCGAATAGAGGTATCGTTGGTTA 45301
 QY 327 AACTGTTGCTAGTAGCTCCCTGCAATTTGGTTCCGCTGAGATTGCCACTGNTCAAAATTCGG 386
 Db 45302 AACTGTTGCTAGTAGCTCCCTGAGGTTGGTTCTGCTGAGATTGCCACTGNTCAAAATTCGT 45361
 QY 387 ACAGATTCCATGTTAGCGCCGCTGCCAGCTGGTTGTTCTGTTGTTATGATCCACATAATAA 446
 Db 45362 ACCAATTCATGTTAGCGCCGCTGCCAGCTGGTTGTTCTGTTGTTATGATCCACATAATAA 45421
 QY 447 ACTCCTCTAGCATCATCACTCTCTTCCCACTGGAGTAAAGGTTGTTGTTCTCTCC 506
 Db 45422 GCTCTTCCAGCATCATCACTCTCTTCCCACTGGAGTAAAGGTTGTTGTTCTCTCC 45481
 QY 507 CATGTGTGTTGTTGAGCATTTATGANCACATAATACGCTCTTACCATGAGGATCTTTTCTT 566
 Db 45482 CGTGTGTGTTGTTGAGCATTTATGATCCACATAATAGTTCTACCATGAGGATTTTCTT 45541
 QY 567 TGTTCACACCCCTGATGGCAANGTTCTGTGTGGCATTTCCAGACTGCTGCCGNAC 623
 Db 45542 TGTTCACCA-CCCTGATGGCAAGGTTTCTGTGTGGCATTTCCAGACTGCTGCTATAC 45597

RESULT 11
 AF412122/c 1794 bp mRNA linear VRT 12-MAR-2002
 LOCUS Gallus gallus late domain-interacting protein 2 mRNA, partial
 DEFINITION sequence.
 ACCESSION AF412122.1 GI:19401482
 VERSION AF412122.1
 KEYWORDS Gallus gallus (chicken)
 SOURCE Gallus gallus
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 1 (bases 1 to 1794)
 Kikonyogo, A., Bouamr, F., Vana, M.L., Xiang, Y., Aliyar, A., Carter, C.
 and Leis, J.
 Proteins related to the Nedda family of ubiquitin protein ligases
 interact with the L domain of Rous sarcoma virus and are required
 for gag budding from cells
 Proc. Natl. Acad. Sci. U.S.A. 98 (20), 11199-11204 (2001)
 21457266
 11562473
 PUBLISHED
 REFERENCE 2 (bases 1 to 1794)
 Kikonyogo, A., Bouamr, F., Vana, M.L., Xiang, Y., Aliyar, A., Carter, C.
 and Leis, J.
 Direct Submission
 TITLE Submitted (22-AUG-2001) Microbiology and Immunology, Northwestern
 University, 303 E. Chicago, Chicago, IL 60611, USA
 JOURNAL Location/Qualifiers
 FEATURES
 source
 1. 1794
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /db_xref="taxon:9031"
 misc_feature 100..>1794
 /note="similar to E3 ubiquitin ligase; late domain-interacting protein 2; LDI-2; coding region not yet determined"
 misc_feature 1162..1251
 /note="potential WW 1"
 misc_feature 1255..1347
 /note="potential WW 2"
 misc_feature 1483..1569
 /note="potential WW 3"
 misc_feature 1597..1692

/note="potential WW 4"

ORIGIN

Query Match 69.1%; Score 431.2; DB 5; Length 1794;
 Best Local Similarity 81.8%; Pred. No. 2.5e-107;
 Matches 504; Conservative 0; Mismatches 111; Indels 1; Gaps 1;
 QY 9 TTNTAAGCAATTTGTGGACCCCTTTAGTNCAGATGACTTCCCAATTCGAGGATCTTTGAAT 68
 Db 1736 TCATAGGCAATCTGTGGCCCTTTTATTATTACAGAAGACTTCCAGTACGAGGATCATTTGAAG 1677
 QY 69 GTTGTGTTCTTGTGTATGATCAACAAAGTACTTACACTTACAGAGTATATCTAATT 128
 Db 1676 GTGGTAGTCTTGTATATGATCAACAAAGTATCGTACACCTTCTCTAGTGTATCTGATT 1617
 QY 129 TCCAGAGCTTCTGGCAGGGTCTTCAATCTGTAAAGCTTTGAGTTCCTTGGATCTTCCAC 188
 Db 1616 TCCAGCCCTCTGAGAGAGGTCTCAATTTGTAAACCTTGGTTCGAGGGTCTTCCCAT 1557
 QY 189 TGGGTTGTTTGTGTATGATTACAAAGTAAACCCCTGTCTGTGAATCCACTCTTTT 248
 Db 1556 TGTGTTGTCTTGTGTATGATTACAAATACACCTATCATTTGAATCCACTCTCCTT 1497
 QY 249 TCCAGGCTGGTGGCAAGGTCCATTAAGGTCAATTTCTGCAGCTAATTAAGCCGAA 308
 Db 1496 TCCAGCCCTGTGGCAAGGCCCAAGTGGATCGTTCTCTGCTGACAAATCGAAGCCGAA 1437
 QY 309 TAGAGGTATCTGTGGTTAAACTGTTCATAGCTCCCTGCAATTTGGTTCGCTGAGATTGC 368
 Db 1436 TAGAGGTATCTGTGGTTAAATTTGTCATAGCTCCCTGCAATTTGATTAACCTGAGATGC 1377
 QY 369 CACTGNTCAAAATTCGGACAGATTCCATGTTAGGCGCTGCCACGTTGTTGTTCTGGTG 428
 Db 1376 CACTGCTCAAAGTTCTCTGACTGATTCATTTGTTGGTGGCTGCCATGTTGTTGTTCTGGTG 1317
 QY 429 TTATGATCCACATPAATAAATCTCTTACGATCATCAACTCTTCTTCCCAACCTCGAGGT 488
 Db 1316 TTATGGTCCACATAGTAAACTCTTCCACGATCATCAACTCTTCTTCCCAACCTCGAGGT 1257
 QY 489 AAAGGTTGGTCTCTCCCATGTTGGTGTAGTTCGAGAGCATTAAGCCACATTAATAGCTTA 548
 Db 1256 AAGGGCTGTGGTCTTTCCTCCACGTTGTAGTTCGTTGATTTGATGCAACATAATAGTTCTA 1197
 QY 549 CCATGAGGATCTTTCTTTTCTTCCACCCCTGATGCGCAANGNTTCTGTGTGGCATTTCCC 608
 Db 1196 CCATGAGGATCTTTCTTTTCTTCCCA-CCCTGGTGTAAAGGTTCTGTACCAGCATTAAC 1138
 QY 609 AGACTGCTGCCGNACA 624
 Db 1137 AGTCTGCTGCTTACA 1122

RESULT 12
 BC076842/c

LOCUS BC076842 3501 bp mRNA linear VRT 19-JUL-2004
 DEFINITION Xenopus laevis cDNA clone IMAGE:6859998, partial cds.
 ACCESSION BC076842
 VERSION BC076842.1 GI:50368904
 KEYWORDS
 SOURCE Xenopus laevis (African clawed frog)
 ORGANISM Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 Xenopodinae; Xenopus; Xenopus.
 1 (bases 1 to 3501)
 Klein, S.L., Strausberg, R.L., Wagner, L., Pontius, J., Clifton, S.W.
 and Richardson, P.
 Genetic and genomic tools for Xenopus research: The NIH Xenopus
 initiative
 Dev. Dyn. 225 (4), 384-391 (2002)
 JOURNAL PUBLISHED 12454917
 REFERENCE 2 (bases 1 to 3501)
 STRAUSBERG, R.L., FEINGOLD, E.A., GROUSE, L.H., DERGE, J.G.,

Db 76124 CTTTCCCAACTGGAGCA 76142
|||||

RESULT 14

AC103927

LOCUS

DEFINITION

AC103927

AC103927.4

VERSION

KEYWORDS

SOURCE

- ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

AC103927 135660 bp DNA linear HTG 09-MAR-2004
Mus musculus clone RP23-44A16, WORKING DRAFT SSEQUENCE, 12 unordered
pieces.

AC103927 4 GI:44681626
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
Mus musculus (house mouse)

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Birren,B., Nussbaum,C. and Lander,B.
Mus musculus, clone RP23-44A16
Unpublished

2 (bases 1 to 135660)
Birren,B., Linton,L., Nussbaum,C., Lander,B., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepe,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., Lakocque,K.,
Lanazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheters,R., Meldrum,J.,
Meneus,L., Mihova,T., Mlewa,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rottet,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (01-DEC-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 135660)
Birren,B., Nussbaum,C., Lander,B., Abouelleil,A., Allen,N.,
Anderson,M., Atachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepe,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., DeArrellano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,X., Liu,A., Mabbitt,R., Maclean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meldrum,J., Meneus,L., Mihova,T., Mlewa,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schupack,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (09-MAR-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

On Feb 28, 2004 this sequence version replaced gi:31455749.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR

Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information
Center project name: L17179

Center clone name: 44_A_16
----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731
Consensus quality: 131840 bases at least Q40

Consensus quality: 132850 bases at least Q30
Consensus quality: 133337 bases at least Q20

Insert size: 127000; agarose-fp
Insert size: 134128; sum-of-contigs

Quality coverage: 12.1 in Q20 bases; agarose-fp
Quality coverage: 11.5 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1
* 13935: contig of 13935 bp in length
* 13936 14035: gap of 100 bp

* 14036 15137: contig of 1282 bp in length
* 15138 15417: gap of 100 bp
* 15418 17222: contig of 1805 bp in length

* 17223 17322: gap of 100 bp
* 17323 22104: contig of 4782 bp in length
* 22105 22204: gap of 100 bp

* 22205 31776: contig of 9572 bp in length
* 31777 39882: contig of 8006 bp in length
* 39883 39982: gap of 100 bp

* 39983 65136: contig of 25154 bp in length
* 65137 65236: gap of 100 bp
* 65237 75051: contig of 9815 bp in length

* 75052 75151: gap of 100 bp
* 75152 92728: contig of 17577 bp in length
* 92729 92828: gap of 100 bp

* 92829 115139: contig of 22311 bp in length
* 115140 115239: gap of 100 bp
* 115240 116031: contig of 792 bp in length

* 116032 116131: gap of 100 bp
* 116132 135660: contig of 19529 bp in length.

FEATURES

source

1..135660
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-44A16"
/clone_lib="RP23-44A16" Female Mouse BAC"

misc_feature

1..13935
/note="assembly_fragment"
clone_end:SP6
vector_side:left
14036..15317

misc_feature

/note="assembly_fragment"
15418..17222
/note="assembly_fragment"
17323..22104

misc_feature

/note="assembly_fragment"
22205..31776
/note="assembly_fragment"
31877..39882
/note="assembly_fragment"

```
misc_feature 39983..65136
/note="assembly_fragment"
misc_feature 65237..75051
/note="assembly_fragment"
misc_feature 75152..92728
/note="assembly_fragment"
misc_feature 92829..115139
/note="assembly_fragment"
misc_feature 115240..116031
/note="assembly_fragment"
misc_feature 116132..135660
/note="assembly_fragment
clone_end:17
vector_side:right"

ORIGIN
Query Match 57.08; Score 355.4; DB 2; Length 135660;
Best Local Similarity 90.98; Pred. No. 8.8e-87;
Matches 399; Conservative 0; Mismatches 38; Indels 2; Gaps 2;

QY 51 TTGCGAGGATCTTTGAATGTTGTTGTTCTTGTGTTATGATCAACAAGTACCTTACACCT 110
Db 112063 TTCCAGGATCTTTGAATGTTGTTCTGCTGTTATGATCAACAAGTA-CTAACACCT 112121

QY 111 TCAGGATATATCAATTTCCAGCCTTCTGCGAGGGTCTTCAATCTGTAGCCCTTGA 170
Db 112122 TCCGAGTATACCTAAATTTCCAGCCTTCTGCGAGGGTCTTCAATTTGGTAAAGCCTTGA 112181

QY 171 GTTCTTGATCTTCCCACTGGTGTGTTGTTGTTATGATTCACAAAGTAAACCTGTCT 230
Db 112182 GTTCTTGATCTTCCCACTGGTGTGTTGTTGTTATGATTCACAAAGTAAAGTCTGTCT 112241

QY 231 GTTGAATCCACTCTTTTTCCTCCAGCCTGTTGCGCAAGGTCATTAAGGTCATTTCTGCA 290
Db 112242 GTTGAATCCACTCTTTTTCCTCCAGCCTGTTGCGCAAGGTCATTTCTGCA 112301

QY 291 GCTAACATTCGAAGCCGATAGAGTATCGTTGGTAACTGTTGTCATAGCTCCCTGCAAT 350
Db 112302 GCTAACATTCGAAGCCGATAGAGTATCGTTGGTAACTGTTGTCATAGCTCCCTGCAAT 112360

QY 351 TGGTTCGCTGAGATTCGCCACTGTCCTCAAAATTCGAGACATCCATGAGGCGCTGC 410
Db 112361 TGGTTCGCTGAGATTCGCCACTGTCCTCAAAATTCGAGACATCCATGAGGCGCTGC 112420

QY 411 CAGCTTCTGTTCTGTTGTTATGATCCACATAAATACTCTCTTACCATCATCAACTCTTT 470
Db 112421 CAGCTTCTGTTCTGTTGTTATGATCCACATAAATACTCTCTTACCATCATCAACTCTTT 112480

QY 471 CTTTCCCAACTGGAGTA 489
Db 112481 CTTTCCCAACTGGAGTA 112499

RESULT 15
AC101785/c
LOCUS AC101785 189541 bp DNA linear HTG 12-OCT-2002
DEFINITION Mus musculus clone RP24-136P4, WORKING DRAFT SEQUENCE, 22 unordered
pieces.
ACCESSION AC101785
VERSION AC101785.2 GI:23915573
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 189541)
Birren,B., Linton,B., Nussbaum,C. and Lander,E.
TITLE Mus musculus, clone RP24-136P4
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 189541)
AUTHORS Birren,B., Linton,B., Nussbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepe,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
Lamaze,R., Landers,T., Lenockzy,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marquis,N., Mathews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Rile,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 189541)
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L17398
Center clone name: 136_P_4
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 178591 bases at least Q40
Consensus quality: 184996 bases at least Q30
Consensus quality: 186786 bases at least Q20
Insert size: 157000; agarose-fp
Insert quality: 187441; sum-of-contigs
Quality coverage: 11.2 in Q20 bases; agarose-fp
Quality coverage: 9.4 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 22 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
```

```
Choepe,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
Lamaze,R., Landers,T., Lenockzy,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marquis,N., Mathews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Rile,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 189541)
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L17398
Center clone name: 136_P_4
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 178591 bases at least Q40
Consensus quality: 184996 bases at least Q30
Consensus quality: 186786 bases at least Q20
Insert size: 157000; agarose-fp
Insert quality: 187441; sum-of-contigs
Quality coverage: 11.2 in Q20 bases; agarose-fp
Quality coverage: 9.4 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 22 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
```


* This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 692: contig of 692 bp in length
 693 792: gap of 100 bp
 793 1433: contig of 641 bp in length
 1434 1533: gap of 100 bp
 1534 2667: contig of 1134 bp in length
 2668 2767: gap of 100 bp
 2768 3958: contig of 1191 bp in length
 3959 4059: gap of 100 bp
 4060 4761: contig of 703 bp in length
 4762 4861: gap of 100 bp
 4862 6373: contig of 1512 bp in length
 6374 7488: contig of 1015 bp in length
 7489 7589: gap of 100 bp
 7590 8332: contig of 744 bp in length
 8333 8432: gap of 100 bp
 8433 10101: contig of 1669 bp in length
 10102 10201: gap of 100 bp
 10202 11524: contig of 1323 bp in length
 11525 11625: gap of 100 bp
 11626 13294: contig of 1669 bp in length
 13295 13394: gap of 100 bp
 13395 14364: contig of 971 bp in length
 14365 14464: gap of 100 bp
 14465 16117: contig of 1653 bp in length
 16118 16218: gap of 100 bp
 16219 17411: contig of 1193 bp in length
 17412 17511: gap of 100 bp
 17512 19574: contig of 2064 bp in length
 19575 19674: gap of 100 bp
 19675 21717: contig of 2042 bp in length
 21718 21816: gap of 100 bp
 21817 24076: contig of 2260 bp in length
 24077 24176: gap of 100 bp
 24177 27832: contig of 3656 bp in length
 27833 27932: gap of 100 bp
 27933 48374: contig of 20442 bp in length
 48375 48474: gap of 100 bp
 48475 63552: contig of 15078 bp in length
 63553 63653: gap of 100 bp
 63654 101162: contig of 37510 bp in length
 101163 101262: gap of 100 bp
 101263 189541: contig of 88279 bp in length.

FEATURES

source
 1. .189541
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /clone="RP24-136P4"
 /clone_lib="RPCI-24 Male Mouse BAC"
 1. .692
 /note="assembly_fragment"
 793. .1433
 /note="assembly_fragment"
 1534. .2667
 /note="assembly_fragment"
 2768. .3958
 /note="assembly_fragment"
 4059. .4761
 /note="assembly_fragment"
 4862. .6373
 /note="assembly_fragment"
 6474. .7488
 /note="assembly_fragment"
 7589. .8332
 /note="assembly_fragment"
 8433. .10101
 /note="assembly_fragment"
 10202. .11524
 /note="assembly_fragment"

misc_feature 11625. .13293
 /note="assembly_fragment"
 misc_feature 13394. .14364
 /note="assembly_fragment"
 misc_feature 14465. .16117
 /note="assembly_fragment"
 misc_feature 16218. .17410
 /note="assembly_fragment"
 misc_feature 17511. .19574
 /note="assembly_fragment"
 misc_feature 19675. .21716
 /note="assembly_fragment"
 misc_feature 21817. .24076
 /note="assembly_fragment"
 misc_feature 24177. .27832
 /note="assembly_fragment"
 misc_feature 27933. .48374
 /note="assembly_fragment"
 misc_feature 48475. .63552
 /note="assembly_fragment"
 misc_feature 63653. .101162
 /note="assembly_fragment"
 misc_feature 101263. .189541
 /note="assembly_fragment"

ORIGIN

Query Match 57.0%; Score 355.4; DB 2; Length 189541;
 Best Local Similarity 90.9%; Pred. No. 8.4e-87;
 Matches 399; Conservative 0; Mismatches 38; Indels 2; Gaps 2;
 QY 51 TTGCGAGGATCTTCAATGTTGTTCTTCTGCGAGGGTCTTCTTCAATCTGTAAGCTTGA 170
 Db 29461 TTCCCGAGGATCTTCAATGTTGTTCTTCTGCGAGGGTCTTCTTCAATCTGTAAGCTTGA 170
 QY 111 TCACGAGTATATCTAATTTCCAGCCTTCTGCGAGGGTCTTCTTCAATCTGTAAGCTTGA 170
 Db 29402 TCCCGAGTATACCTAATTTCCAGCCTTCTGCGAGGGTCTTCTTCAATCTGTAAGCTTGA 170
 QY 171 GTTCTTGGATCTTCCCACTGGTGTGTTTCTGTTGTTATGATTCACAAAGTAAACCTGTCT 230
 Db 29342 GTTCTTGGATCTTCCCACTGGTGTGTTTCTGTTGTTATGATTCACAAAGTAAAGTCTGTCT 29283
 QY 231 GTTGAATCCACTCTTTTTCAGCCTGTTGCGCAAGTCCATAGGTCATTTCTGCA 290
 Db 29282 GTTGAATCCACTCTTTTTCAGCCTGTTGCGCAAGTCCATAGGTCATTTCTGCA 29223
 QY 291 GCTAACATTGAAGCGGAATAGAGTATCGTTGGTTAAACTGTTGCATAGCTCCCTGCAAT 350
 Db 29222 GCTAACATTGAAGCGGAATAGAGTATCGTTGGTTAAACTGTTGCATAGCTCCCTGCAAT 29164
 QY 351 TGGTTCCGCTGAGATTGCCACTGNTCAAAATTCGGACAGATCCATGTTAGCGCGTGC 410
 Db 29163 TGGTTCCGCTGAGATTGCCACTGNTCAAAATTCGGACAGATCCATGTTAGCGCGTGC 29104
 QY 411 CACGTTGTTGTTGTTGTTGTTATGATCCACATAATAAATCTCTTACGATCATCAACTCTT 470
 Db 29103 CAGGTTCTTGTCTGTTGTTGTTATGATCCACATAATAAATCTCTTACGATCATCAACTCTT 29044
 QY 471 CTTTCCCAACTCGAGGTA 489
 Db 29043 CTTTCCCAACTCGAGGTA 29025

Search completed: June 30, 2005, 05:30:24
 Job time : 3093 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2005, 03:18:24 ; Search time 471 Seconds
(without alignments)
7842.715 Million cell updates/sec

Title: US-10-618-408-3
Perfect score: 624
Sequence: 1 ccgcgcgtntaagcattt.....tccagactgctgcgcnaca 624

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	616	98.7	624	13	AdS00022 Human WWP
2	566.2	90.7	2052	2	Aat95697 Novel hum
3	566.2	90.7	2052	4	Aad10504 Human ubi
4	566.2	90.7	2052	6	Abz34944 Human gen
5	566.2	90.7	2052	8	Aad50455 Nedd-4-1i
6	566.2	90.7	2052	10	Adb49239 Novel hum
7	566.2	90.7	2731	10	Adc26288 Human NOV
8	566.2	90.7	2793	13	AdS00020 Human WWP
9	566.2	90.7	2827	4	Aac90422 Human clo
10	566.2	90.7	3382	6	Abaz93745 Human sig
11	563.6	90.3	943	13	Acn38827 Tumour-as
12	552.2	88.5	3495	4	Aac90423 Human hom
13	456.8	73.2	2305	12	Adp84403 Human bre
14	246.6	39.5	2217	5	Aas80664 DNA encod
15	246.6	39.5	2377	5	Abaz82679 Atrophin-
16	246.6	39.5	2377	8	Aad50454 Atrophin-
17	246.6	39.5	2377	8	Acc45999 Human atr
18	246.6	39.5	2377	10	Adb98693 Human atr
19	246.6	39.5	2377	10	Ade84909 Farnesyl
20	246.6	39.5	2377	10	Ade82489 Human DNA

C 21	246.6	39.5	2559	2	AAZ09236	Aaz09236 Human E3
C 22	246.6	39.5	2678	12	ADE77069	Ade77069 Human CDN
C 23	246.6	39.5	3410	12	ADQ85919	Adq85919 Human tum
C 24	246.6	39.5	3410	12	ADQ84071	Adq84071 Human tum
C 25	246.6	39.5	3745	8	ACF34485	Acf34485 Gene enco
C 26	246.6	39.5	5372	2	AAZ09235	Aaz09235 Human E3
C 27	210	33.7	3288	13	ACN40036	Acn40036 Tumour-as
C 28	203.6	32.6	3475	4	AAD10511	Aad10511 Human ubi
C 29	203.6	32.6	3475	8	AAD50456	Aad50456 Nedd-4-1i
C 30	203.6	32.6	3475	12	ADQ19347	Adq19347 Human sof
C 31	203.6	32.6	3476	2	AAT95698	Aat95698 Novel hum
C 32	203.6	32.6	3476	10	ADB49241	Adb49241 Novel hum
C 33	168.8	27.1	4573	12	ADQ23594	Adq23594 Human sof
C 34	168.8	27.1	4573	12	ADQ23375	Adq23375 Human sof
C 35	135.2	21.7	4307	4	ABL06079	Ab106079 Drosophil
C 36	135.2	21.7	4390	4	ABL05629	Ab105629 Drosophil
C 37	111.8	17.9	2351	2	AAZ30406	Aaz30406 DNA encod
C 38	111.8	17.9	2351	10	ADB47742	Adb47742 Novel hum
C 39	111.8	17.9	2351	12	ADJ55297	Adj55297 Novel hum
C 40	95.6	15.3	3663	12	ADQ64872	Adq64872 Novel hum
C 41	91.2	14.6	8354	4	ABL06078	Ab106078 Drosophil
C 42	91.2	14.6	8653	4	ABL05628	Ab105628 Drosophil
C 43	88.4	14.2	349	5	AAS80663	Aas80663 DNA encod
C 44	59.2	9.5	1863	5	AAS80665	Aas80665 DNA encod
C 45	56	9.0	2661	10	ADB53839	Adb53839 Primary r

ALIGNMENTS

RESULT 1
ADS00022
ID ADS00022 standard; DNA; 624 BP.

AC ADS00022;

DT 16-DEC-2004 (first entry)

DE Human WWP1 antisense fragment.

KW ds; gene; apoptosis-related disease; human; WWP1; apoptosis; cancer.

OS Homo sapiens.

PN US2004191220-A1.

PD 30-SEP-2004.

PF 11-JUL-2003; 2003US-00618408.

PR 11-JUL-2002; 2002US-0395358P.

XX (EINA/) EINAT P.

PA (DEIS/) DEISS L.

XX (MAYA/) MAYA R.

PI Einat P, Deiss L, Maya R;

XX WPI; 2004-698655/68.

DR Use of an inhibitor of the WWP1 polypeptide for treating an apoptosis-related disease in a subject such as cancer, and for preparing a medicament for treating the apoptosis-related disease.

XX Claim 4; SEQ ID NO 3; 32pp; English.

PS The invention relates to a method of treating an apoptosis-related disease in a subject which comprises administering to the subject a therapeutic amount of an inhibitor of the WWP1 polypeptide to inhibit WWP1 so as to thus treat the subject. The method is useful for treating an apoptosis-related disease, i.e. cancer. Inhibitor of human WWP1 polypeptide is useful in preparing a medicament for treating cancer. The present sequence represents the human WWP1 antisense fragment.

XX SQ Sequence 624 BP; 132 A; 150 C; 130 G; 204 T; 0 U; 8 Other;

Query Match 98.7%; Score 616; DB 13; Length 624;
 Best Local Similarity 100.0%; Pred. No. 2.9e-184;
 Matches 624; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGGTTTAAAGCAATTTGAGACCCCTTTAGTNCAGATGACTTCCATTGCGAGGAT 60
 DB 1 CGCGCGGTTTAAAGCAATTTGAGACCCCTTTAGTNCAGATGACTTCCATTGCGAGGAT 60

QY 61 CTTTGAATGTTGTTCTTCTGTTATGATCAACAAGTACCTTTACAGAGTAT 120
 DB 61 CTTTGAATGTTGTTCTTCTGTTATGATCAACAAGTACCTTTACAGAGTAT 120

QY 121 ATCTAAATTTCCAGCTTCTGCGAGGGGTTCTTCAATCTGTAAAGCTTGTGAT 180
 DB 121 ATCTAAATTTCCAGCTTCTGCGAGGGGTTCTTCAATCTGTAAAGCTTGTGAT 180

QY 181 CTTCCACTGGGTTGTTTCTGTTATGATCAACAAGTAAACCTGCTGTGTAATCCA 240
 DB 181 CTTCCACTGGGTTGTTTCTGTTATGATCAACAAGTAAACCTGCTGTGTAATCCA 240

QY 241 CTTCTTTTCCAGCTGCTGGGAAAGTCCATTAAGGGTCAATTTCTGCGAGTAAATG 300
 DB 241 CTTCTTTTCCAGCTGCTGGGAAAGTCCATTAAGGGTCAATTTCTGCGAGTAAATG 300

QY 301 AAGCGAATAGAGTATGTTGTTTAAACTGTTGCATAGCTCCCTGCAATTTGGTTCGCT 360
 DB 301 AAGCGAATAGAGTATGTTGTTTAAACTGTTGCATAGCTCCCTGCAATTTGGTTCGCT 360

QY 361 GAGATGCCACTGNTCAAAATTTGCGACAGATTCATGTTGAGCGCTGCGACGTTGTTG 420
 DB 361 GAGATGCCACTGNTCAAAATTTGCGACAGATTCATGTTGAGCGCTGCGACGTTGTTG 420

QY 421 TTCTGGTTTATGATACACATAATAAACTCTCTACGATCATGACCTCTCTTCTTCCCAAC 480
 DB 421 TTCTGGTTTATGATACACATAATAAACTCTCTACGATCATGACCTCTCTTCTTCCCAAC 480

QY 481 CTGGAGTAAAGGTTGTGGTCTCTCCCATGTTGGTATGTTGAGCATTTATGANCACATAAT 540
 DB 481 CTGGAGTAAAGGTTGTGGTCTCTCCCATGTTGGTATGTTGAGCATTTATGANCACATAAT 540

QY 541 ACGCTCTACCATGAGGATCTTTTCTTTGTTCCCAACCCCTGATGGCAANGTCTGTGTTG 600
 DB 541 ACGCTCTACCATGAGGATCTTTTCTTTGTTCCCAACCCCTGATGGCAANGTCTGTGTTG 600

QY 601 GCATTTCCAGACTGCTGCCGNACA 624
 DB 601 GCATTTCCAGACTGCTGCCGNACA 624

RESULT 2
 ID AAT95697/c
 XX AAT95697 standard; DNA; 2052 BP.
 AC AAT95697;

XX 23-APR-1998 (first entry)
 XX Novel human gene, designated WWP1.

XX Peptide recognition unit; YAP WW domain binding protein; WBP-1; WBP-2;
 KW WW domain; cell signalling; growth regulation; cytoskeleton organisation;
 KW targeted drug screening; modulator; WW domain interaction; ss.
 XX Homo sapiens.
 OS
 XX WO9737223-A1.
 PN
 XX 09-OCT-1997.
 PD
 XX 03-APR-1997; 97WO-US0005547.

XX 03-APR-1996; 96US-00630916.
 PR (CYTO-) CYTOGEN CORP.
 PA (UYNC-) UNIV NORTH CAROLINA.
 XX Pirozzi G, Kay BK, Fowlkes DM;
 XX WPI; 1997-503234/46.
 DR P-PSDB; AAW36794.

XX Identifying cell signalling and growth regulatory polypeptides by
 PT reaction with multivalent recognition complex - polypeptides are useful
 PT in targetted drug selection.

XX Claim 68; Fig 16 A-B; 220pp; English.

XX The present DNA sequence encodes a novel protein WWP1. The WWP1 gene was
 CC identified and isolated from human bone marrow and brain CDNA libraries,
 CC using peptides AAW38103-05. These peptide recognition units are based on
 CC the sequences of WW domain binding domains of the YAP WW domain binding
 CC proteins WBP-1 and WBP-2. The WW domain is a small functional domain
 CC found in a large number of proteins from a variety of species including
 CC humans, nematodes and yeast. Its name is derived from the observation
 CC that two tryptophan residues, one in the amino terminal portion of the WW
 CC domain and one in the carboxyl terminal portion, are conserved. Most
 CC proteins containing WW domains have a function involving cell signalling
 CC and growth regulation or the organisation of the cytoskeleton.

CC Polypeptides containing a WW domain are identified by treating a
 CC multivalent recognition unit complex that has selective binding affinity
 CC for a WW domain, with many polypeptides and identifying those with
 CC selective affinity for the complex. Proteins containing WW domains are
 CC used for targeted drug screening, i.e. to identify potential modulators
 CC of specific WW domain interactions. The valency of the recognition unit
 CC is important in determining specificity of interaction with WW domains.
 CC In multivalent form specificity is relaxed, but not lost, so proteins
 CC containing WW domains similar, but not identical, to the sequence of the
 CC peptides' target WW can be detected, including new polypeptides

XX SQ Sequence 2052 BP; 636 A; 405 C; 447 G; 564 T; 0 U; 0 Other;

Query Match 90.7%; Score 566.2; DB 2; Length 2052;
 Best Local Similarity 97.3%; Pred. No. 3e-168;
 Matches 582; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

QY 27 CCCTTTAGTNCAGATGACTTCCCATTCGAGAGATCTTTGAATGTTGTTCTTCTGTTA 86
 DB 1044 CCTTTAGTTACAGATGACTTCCCATTCGAGAGATCTTTGAATGTTGTTCTTCTGTTA 985

QY 87 TGATCAACAAGTACCTTACACCTTTCAGAGTATATCTAAATTTCCAGCTTCTGCGAGG 146
 DB 984 TGATCAACAAGTACCTTACACCTTTCAGAGTATATCTAAATTTCCAGCTTCTGCGAGG 925

QY 147 GGTTCCTTCTGTTAAGCTTGAAGTCTTCCCATTCGAGTATCTTCCCTGCGGTTGTTTGTGTTA 206
 DB 924 GGTTCCTTCTGTTAAGCTTGAAGTCTTCCCATTCGAGTATCTTCCCTGCGGTTGTTTGTGTTA 865

QY 207 TGATTCACAAGTAAACCCCTGCTGTAATCCACTCTTTTCCAGCTGTTGTCGCAAA 266
 DB 864 TGATTCACAAGTAAACCCCTGCTGTAATCCACTCTTTTCCAGCTGTTGTCGCAAA 805

QY 267 GGTTCATAGGGTCAATTTTCTGAGCTAACATTTGAAGCCGAATAGAGGTATCGTTGGTTA 326
 DB 804 GGTTCATAGGGTCAATTTTCTGAGCTAACATTTGAAGCCGAATAGAGGTATCGTTGGTTA 745

QY 327 RACTGTTGCATAGCTCCCTGCAATTTGGTTCGCTGAGATTCGCCACTGTTCAAAATTCGG 386
 DB 744 AACTGTTGCATAGCTCCCTGCAATTTGGTTCGCTGAGATTCGCCACTGTTCAAAATTCGG 685

QY 387 ACAGATTCATGGTAGGCGCTGCCACGTTGTTGTTCTGCGTGTATGATCCACATAATAA 446
 DB 684 ACAGATTCATGGTAGGCGCTGCCACGTTGTTGTTCTGCGTGTATGATCCACATAATAA 625

QY	447	ACTCCTCTACGATCATCAACTCTTCTTCCCAACCTGGAGGTAAAGGTTGTGTCTCTCC	506
Db	624	ACTCTTCTACGATCATCAACTCTTCTTCCCAACCTGGAGGTAAAGGTTGTGTCTCTCC	565
QY	507	CATGTGTGTAGTTCGAGCATATGACCACATAATACGCTCTACCATGAGGATCTTTTCTT	566
Db	564	CATGTGTGTAGTTCGAGCATATGATCCACATATAGGTTCTACCATGAGGATCTTTTCTT	505
QY	567	TGTTCCCAACCCCTGATGGCAANGNTTCTGTCTTGGCAATCCCAAGACTGCTGCCGNACA	624
Db	504	TGTTCCCA-CCCTGATGGCAAGGTTCTGTGTGGCAATCCCAAGACTGCTGCCGTACA	448
RESULT 3			
AAD10504/c			
ID	AAD10504 standard; DNA; 2052 BP.		
XX			
AC	AAD10504;		
XX			
DT	24-SEP-2001 (first entry)		
XX			
DE	Human ubiquitin protein ligase WWP1 DNA.		
XX			
KW	Human; ubiquitin protein ligase; WWP1; antitumour; antiinflammatory;		
KW	therapy; infection; inflammation; tumour; ds.		
XX			
OS	Homo sapiens.		
PH	Key		
FT	CDS		
FT	2..2050		
FT	/*tag= a		
FT	/product= "Human WWP1 protein"		
FT	/note= "CDS does not include start and stop codon"		
FT	/partial		
XX			
FN	US6258601-B1.		
XX			
PD	10-JUL-2001.		
XX			
PF	07-SEP-2000; 2000US-00657481.		
XX			
PR	07-SEP-2000; 2000US-00657481.		
XX			
PA	(ISIS-) ISIS PHARM INC.		
XX			
PI	Monia BP, Cowesert LM;		
XX			
DR	WPI; 2001-450370/48.		
DR	P-PSDB; AAE05494.		
XX			
PT	Antisense compounds capable of modulating expression of ubiquitin protein		
PT	ligases WWP1 and WWP2, useful for diagnosis, prophylaxis and treatment of		
PT	diseases e.g. infection, inflammation or tumors.		
XX			
PS	Claim 1; Col 53-58; 47pp; English.		
XX			
CC	The present invention relates to compounds, particularly antisense		
CC	oligonucleotides, which are targeted to nucleic acids encoding ubiquitin		
CC	protein ligases WWP1 and WWP2. The antisense oligonucleotides modulate		
CC	the expression of WWP1 and WWP2. The antisense oligonucleotides are		
CC	useful for inhibiting the expression of ubiquitin protein ligases WWP1		
CC	and WWP2 in cells or tissues in vitro. The oligonucleotides are useful		
CC	for diagnosing, treating diseases associated with the expression of		
CC	ubiquitin protein ligases WWP1 and WWP2 and for prophylaxis e.g. to		
CC	prevent or delay infection, inflammation or tumour formation and as a		
CC	research reagent. The present sequence is a DNA encoding human ubiquitin		
CC	protein ligase WWP1		
XX			
SQ	Sequence 2052 BP; 636 A; 405 C; 447 G; 564 T; 0 U; 0 Other;		
Query Match 90.7%; Score 566.2; DB 4; Length 2052;			
Best Local Similarity 97.3%; Pred. No. 3e-168;			
Matches 582; Conservative 0; Mismatches 15; Indels 1; Gaps 1;			

QY	27	CCCTTAGTNCAGATGACATCCCATTTGCCAGGATCTTTGAATGTTGTTCTTGTGTTA	86
Db	1044	CTTTTAGTTACAGATGACATCCCATTTGCCAGGATCTTTGAATGTTGTTCTTGTGTTA	985
QY	87	TGATCAACAAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGCCCTCTGGCAGG	146
Db	984	TGATCAACAAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGCCCTCTGGCAGG	925
QY	147	GGTTCTTCATCTTGTAAGCCTTGAGTCTTTGGATCTTCCACATGGGTTGTTTGTGTTA	206
Db	924	GGTTCTTCATCTTGTAAGCCTTGAGTCTTCCACATGGGTTGTTTGTGTTA	865
QY	207	TGATTCACAAAGTAAACCCCTGCTGTGTAATCCACTCTTTTCCAGCCCTGGTGGCAA	266
Db	864	TGATTCACAAAGTAAACCCCTGCTGTGTAATCCACTCTTTTCCAGCCCTGGTGGCAA	805
QY	267	GGTCCATAAGGGTCAATTTCTGCAGCTAACATTTGAAGCCGAATAGAGGTATCGTTGGTTA	326
Db	804	GGTCCATAAGGGTCAATTTCTGCAGCTAACATTTGAAGCCGAATAGAGGTATCGTTGGTTA	745
QY	327	AACTGTTGCATAGCTCCCTGCAATTTGGTTCGCTGAGATTGCCACTGNTCAAATTCGG	386
Db	744	AACTGTTGCATAGCTCCCTGCAATTTGGTTCGCTGAGATTGCCACTGTTCAAATTCGG	685
QY	387	ACAGATTCCATGATAGGCGCTGCCAGTGTGTTGTTCTGGTGTATGATCCACATAATA	446
Db	684	ACAGATTCCATGATAGGCGCTGCCAGTGTGTTGTTCTGGTGTATGATCCACATAATA	625
QY	447	ACTCCTCTACGATCATCAACTCTTTTCCCAACCTGGAGGTAAAGTGTGTCTCTCC	506
Db	624	ACTCCTCTACGATCATCAACTCTTTTCCCAACCTGGAGGTAAAGTGTGTCTCTCC	565
QY	507	CATGTGTAGTTCGAGCATTTATGANCACACATATAGCTCTACCATGAGGATCTTTTCTT	566
Db	564	CATGTGTAGTTCGAGCATTTATGATCCACATATAGTTCATCCATGAGGATCTTTTCTT	505
QY	567	TGTTCCCAACCCCTGATGGCAANGTCTGTGTTGGCAATCCAGACTGCTGCCGNACA	624
Db	504	TGTTCCCA-CCCTGATGGCAAGTTCGTGTGGCAATCCAGACTGCTGCCGTACA	448
RESULT 4			
ABZ34944/c			
ID	ABZ34944 standard; cDNA; 2052 BP.		
XX			
AC	ABZ34944;		
XX			
DT	05-FEB-2003 (first entry)		
XX			
DE	Human gene expression profile polynucleotide SEQ ID NO 56.		
XX			
KW	Human; artery; endothelium; umbilical; vein; aorta; pulmonary artery;		
KW	bronchial epithelium; prostate; muscle; lung fibroblast; osteoblast;		
KW	tumour; microarray; genome mapping; antiviral; antifungal;		
KW	gene expression; gene; ss.		
OS	Homo sapiens.		
XX			
FN	WO200274979-A2.		
XX			
PD	26-SEP-2002.		
XX			
PF	20-MAR-2002; 2002WO-US008456.		
XX			
PR	20-MAR-2001; 2001US-0276947P.		
XX			
PA	(ORTH) ORTHO CLINICAL DIAGNOSTICS INC.		
XX			
PI	Wan J, Wang Y;		
XX			
DR	WPI; 2002-740862/80.		
XX			

PT New gene expression profile generated from primary, endothelial,
 PT epithelial, and muscle cell types, useful for identifying disease
 PT pathologies involving alterations of gene expression, e.g. cancer.
 XX
 PS
 XX Claim 3; Page 256; 850pp; English.

CC The invention relates to a gene expression profile comprising one or more
 CC genes (ABZ34889-ABZ35692) and generated from a cell type. The cell type
 CC is a coronary artery endothelium, umbilical artery or vein endothelium,
 CC aortic endothelium, dermal microvascular endothelium, pulmonary artery
 CC endothelium, myometrium microvascular endothelium, keratinocyte
 CC epithelium, bronchial epithelium, mammary epithelium, prostate
 CC epithelium, renal cortical epithelium, renal proximal tubule epithelium,
 CC small airway epithelium, renal epithelium, umbilical artery smooth
 CC muscle, neonatal dermal fibroblast, pulmonary artery smooth muscle,
 CC dermal fibroblast, neural progenitor cells, skeletal muscle, astrocytes,
 CC aortic smooth muscle, mesangial cells, coronary artery smooth muscle,
 CC bronchial smooth muscle, uterine smooth muscle, lung fibroblast,
 CC osteoblasts or prostate stromal cell. The gene expression profile is used
 CC for determining the level of RNA expression for a sample, determining the
 CC phenotype of a cell and distinguishing cell types. The gene or a protein
 CC expression profile is useful in identifying disease pathologies involving
 CC alterations of gene expression. The assessment of expression profiles may
 CC provide meaningful information with respect to tumour type and stage,
 CC treatment methods, and prognosis. The gene or protein expression profile
 CC may also be used for creating microarrays. The microarray is useful for
 CC genetic and physical mapping of genomes, DNA sequencing, genetic or
 CC medical diagnosis, genotyping of organisms, confirming cell or tissue
 CC identifications and in identifying promising antibiotics, antiviral or
 CC antifungal agents

XX Sequence 2052 BP; 636 A; 405 C; 447 G; 564 T; 0 U; 0 Other;

Query Match 90.7%; Score 566.2; DB 6; Length 2052;
 Best Local Similarity 97.3%; Pred. No. 3e-168;
 Matches 582; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

QY 27 CCCTTAGTACGATGACCTCCATGCGAGGATCTTTGAATGTTGTTCTGTGTTA 86
 DB 1044 CCTTAGTACGATGACCTCCATGCGAGGATCTTTGAATGTTGTTCTGTGTTA 985
 QY 87 TGATCAACAAAGTACCTTACACCTTACGAGTATATCTAATTTCCAGCCTTCCGAGG 146
 DB 984 TGATCAACAAAGTACCTTACACCTTACGAGTATATCTAATTTCCAGCCTTCCGAGG 925
 QY 147 GGTTCCTTCATCTGTAAGCCTTCAGTCTTGGATCTTCCACCTGGGTTGTTTGTGTTA 206
 DB 924 GGTTCCTTCATCTGTAAGCCTTCAGTCTTGGATCTTCCACCTGGGTTGTTTGTGTTA 865
 QY 207 TGATTCACAAAGTAAACCCCTGCTGTTGATCCACTCTTTTCCAGCCTGTTGGCAA 266
 DB 864 TGATTCACAAAGTAAACCCCTGCTGTTGATCCACTCTTTTCCAGCCTGTTGGCAA 805
 QY 267 GGTCCATAGGGTCATTTCTGACGTAAACATTTGAAGCCGATAGAGGTATCGTTGGTTA 326
 DB 804 GGTCCATAGGGTCATTTCTGACGTAAACATTTGAAGCCGATAGAGGTATCGTTGGTTA 745
 QY 327 AACTGTTGCACTGCTCCGCAATTCGTTCCGCTGAGATGCGACATGNTCAAAATTCGG 386
 DB 744 AACTGTTGCACTGCTCCGCAATTCGTTCCGCTGAGATGCGACATGNTCAAAATTCGG 685
 QY 387 ACAGATTCCATGTTAGGCGCTCCACGTTGTTGTTCTGTTGTTATGATCCACATATAA 446
 DB 684 ACAGATTCCATGTTAGGCGCTCCACGTTGTTGTTCTGTTGTTATGATCCACATATAA 625
 QY 447 ACTCCTCTACGATCATCACTCTTCTTCCCAACCTGGAGGTAAGGTTGTGCTCTCC 506
 DB 624 ACTCCTCTACGATCATCACTCTTCTTCCCAACCTGGAGGTAAGGTTGTGCTCTCC 565
 QY 507 CATGCTGTAGTTCGAGCATTTATGANCACATATAGCTCTACCATGAGGATCTTTTCTT 566
 DB 564 CATGCTGTAGTTCGAGCATTTATGATCCACATATAGGTTCTACCATGAGGATCTTTTCTT 505

QY 567 TGTTCACCCCTGATGGCAAGNTTCTGTGTCATTTCCAGACTGCTGCCGACA 624
 DB 504 TGTTCACA-CCCTGATGGCAAGGTTCTGTGTCATTTCCAGACTGCTGCCGACA 448

RESULT 5

AAD50455/c
 ID AAD50455 standard; DNA; 2052 BP.
 AC AAD50455;
 XX 24-MAR-2003 (first entry)
 XX Nedd-4-like ubiquitin-protein ligase WWP1 DNA.
 DE Viral infection; lymphosarcoma; human immunodeficiency virus; hepatitis;
 DE poliomyelitis; HIV; measles; protein therapy; enzyme; gene;
 KW Nedd-4-like ubiquitin-protein ligase; ds.
 KW Unidentified.
 OS
 XX Key Location/Qualifiers
 FT CDS 2..2052
 FT /tag= a
 FT /product= "Nedd-4-like ubiquitin-protein ligase WWP1"
 FT /transl_except= (pos:2051..2052, aa:Ser)
 FT /notes= "CDS does not include start and stop codon"
 FT /partial
 PN WO200209549-A2.
 XX 14-NOV-2002.
 XX 12-MAR-2002; 2002WO-1B002106.
 XX 12-MAR-2001; 2001US-0275224P.
 PR 31-JUL-2001; 2001US-0308958P.
 PR 07-DEC-2001; 2001US-0340170P.
 XX (PROT-) PROTEOLOGICS LTD.
 XX Greener T, Moskowitz H, Reiss Y, Alroy I;
 WPI; 2003-111976/10.
 DR P-PSDB; AAE32721.
 XX New protein complex comprising HECT-RC1, viral maturation scaffolding
 PT protein (VMSF), and/or HIV gag protein, useful for treating viral
 PT infections, such lymphosarcoma, HIV, hepatitis, poliomyelitis, measles,
 PT or Ebola.
 XX Disclosure; Fig 3; 150pp; English.
 XX The invention relates to a method for modulation of viral maturation. The
 CC invention also provides an isolated protein complex comprising a HECT-
 CC RCL1 selected from HECT-WW, HECT-RC1, Gag protein, Gag late domain, P13,
 CC actin, myosin, Hsp60, Hsp90, STAM1, STAM2, VHS-UTM,
 CC GTPase, E2 enzyme, tsq101, cullin, HEC1, HEC2, HEC3, Nedd4 -like
 CC protein or clathrin. The complexes, proteins, antibodies and methods are
 CC useful for treating viral infections, such as lymphosarcoma, human
 CC immunodeficiency virus (HIV), hepatitis, poliomyelitis, measles, or Ebola
 CC and for inhibiting budding in a subject. They are also useful in
 CC diagnostic assays for determining whether a cell is infected with a virus
 CC and for characterising the nature, progression and/or infectivity of the
 CC infection. The invention is also useful in protein therapy. The present
 CC sequence is Nedd-4-like ubiquitin-protein ligase DNA used to illustrate
 CC the method of the invention
 XX Sequence 2052 BP; 636 A; 405 C; 447 G; 564 T; 0 U; 0 Other;

Query Match 90.7%; Score 566.2; DB 8; Length 2052;
 Best Local Similarity 97.3%; Pred. No. 3e-168;
 Matches 582; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

QY 27 CCCTTTAGTNCAGATGACTTCCCATTCGAGGATCTTTGAATGTTGTTGTTCTTGTGTTA 86
 DB 1044 CCTTTAGTTACAGATGACTTCCCATTCGAGGATCTTTGAATGTTGTTGTTCTTGTGTTA 985
 QY 87 TGATCAACAAGTACCTTACACCTTCAGAGTATATCTAATTTCCAGCCCTTCGGCAGG 146
 DB 984 TGATCAACAAGTACCTTACACCTTCAGAGTATATCTAATTTCCAGCCCTTCGGCAGG 925
 QY 147 GGTTCCTTCATCTGTAAAGCCTTGTAGTCTTCCCACTTCCCACTGGTGTGTTGTTGTTA 206
 DB 924 GGTTCCTTCATCTGTAAAGCCTTGTAGTCTTCCCACTTCCCACTGGTGTGTTGTTGTTA 865
 QY 207 TGATTCACAAAGTAAACCCCTGTCTGTGTAATCCACTCTTTTCCAGCCCTTCGGCAAA 266
 DB 864 TGATTCACAAAGTAAACCCCTGTCTGTGTAATCCACTCTTTTCCAGCCCTTCGGCAAA 805
 QY 267 GGTCCATAAGGTCATTTCTGACGCTAACATTAAGCCGAATAGAGGTATCGTTGTTA 326
 DB 804 GGTCCATAAGGTCATTTCTGACGCTAACATTAAGCCGAATAGAGGTATCGTTGTTA 745
 QY 327 AACTGTTGCAATAGCTCCCTGCAATTTGTTTCGCTGAGATGCCACTGNTCAAAATTTCGG 386
 DB 744 AACTGTTGCAATAGCTCCCTGCAATTTGTTTCGCTGAGATGCCACTGNTCAAAATTTCGG 685
 QY 387 ACAGATTCATGAGTGGCCGCTGCCAGCTTTGTTGTTGTTGTTGTTGTTGTTGTTA 446
 DB 684 ACAGATTCATGAGTGGCCGCTGCCAGCTTTGTTGTTGTTGTTGTTGTTGTTGTTA 625
 QY 447 ACTCTCTACGATCATCACTCTTCTTCCCACTGAGGTAAAGTTGTTGTTCTCTCC 506
 DB 624 ACTCTCTACGATCATCACTCTTCTTCCCACTGAGGTAAAGTTGTTGTTCTCTCC 565
 QY 507 CATGTGTTAGTTCGAGCATTAATGACCAATTAAGCCGAATAGAGGTATCGTTGTTA 566
 DB 564 CATGTGTTAGTTCGAGCATTAATGACCAATTAAGCCGAATAGAGGTATCGTTGTTA 505
 QY 567 TGTTCACACCCCTGATGGCAANGNTTCTGTGTTGGCAATTCACAGCTGCTGCCGNACA 624
 DB 504 TGTTCACCA-CCCTGATGGCAANGNTTCTGTGTTGGCAATTCACAGCTGCTGCCGNACA 448

RESULT 6
 ADB49239/c
 ID ADB49239 standard; DNA; 2052 BP.
 XX AC ADB49239;
 XX DT 04-DEC-2003 (first entry)
 XX DE Novel human WWP (WW binding protein) 1 DNA.
 XX KW WW domain; drug candidate screening; drug discovery; drug modification;
 KW drug refinement; immunogen; WWP1; WW binding protein 1; gene; ds; human.
 XX OS Homo sapiens.
 XX PN US2003077577-A1.
 XX PD 24-APR-2003.
 XX XX 28-JUN-2002; 2002US-00185050.
 XX XX 03-APR-1996; 96US-00630916.
 PR 03-APR-1997; 97US-00826516.
 XX (PIROZZI/ PIROZZI G.
 PA (KAYB/) KAY B K.
 PA (FOWL/) FOWLKES D M.
 XX Pirozzi G, Kay BK, Fowlkes DM;
 XX WPI; 2003-635075/60.

DR P-PSDB; ADB49240.
 XX Novel purified polypeptide comprising WW domain, useful for drug
 PT discovery, modification and refinement, for discovering polypeptides
 PT involved in pharmacological activities, or as an immunogen to generate
 PT antibodies.
 XX Claim 68; Fig 16; 133pp; English.
 CC The invention describes a purified polypeptide (I) comprising a WW domain
 CC which has a sequence (S1) selected from 11 sequences fully defined in the
 CC specification, a sequence (S2) selected from 48 sequences fully defined
 CC in the specification or a sequence (S3) comprising 683, 906, 224 or 725
 CC amino acids fully defined in the specification. (I) is useful for
 CC screening a potential drug candidate, by allowing (I) to come into
 CC contact with at least one recognition unit having a selective affinity
 CC for the WW domain in (I), in the presence of an amount of a potential
 CC drug candidate, such that (I) and the recognition unit are capable of
 CC interacting when brought into contact with one another in the absence of
 CC the drug candidate, and determining the effect, if any, of the presence
 CC of the amount of the drug candidate on the interaction of (I) with the
 CC recognition unit. (I) is useful for drug discovery, modification and
 CC refinement, for discovering polypeptides involved in pharmacological
 CC activities, or as an immunogen to generate antibodies. This sequence
 CC encodes a novel human WW binding protein WWP1.
 XX SQ Sequence 2052 BP; 636 A; 405 C; 447 G; 564 T; 0 U; 0 Other;
 Query Match 90.7%; Score 566.2; DB 10; Length 2052;
 Best Local Similarity 97.3%; Pred. No. 3e-168;
 Matches 582; Conservative 0; Mismatches 15; Indels 1; Gaps 1;
 QY 27 CCCTTTAGTNCAGATGACTTCCCATTCGAGGATCTTTGAATGTTGTTGTTCTTGTGTTA 86
 DB 1044 CCTTTAGTTACAGATGACTTCCCATTCGAGGATCTTTGAATGTTGTTGTTCTTGTGTTA 985
 QY 87 TGATCAACAAGTACCTTACACCTTCAGAGTATATCTAATTTCCAGCCCTTCGGCAGG 146
 DB 984 TGATCAACAAGTACCTTACACCTTCAGAGTATATCTAATTTCCAGCCCTTCGGCAGG 925
 QY 147 GGTTCCTTCATCTGTAAAGCCTTGTAGTCTTCCCACTTCCCACTGGTGTGTTGTTGTTA 206
 DB 924 GGTTCCTTCATCTGTAAAGCCTTGTAGTCTTCCCACTTCCCACTGGTGTGTTGTTGTTA 865
 QY 207 TGATTCACAAAGTAAACCCCTGTCTGTGTAATCCACTCTTTTCCAGCCCTTCGGCAAA 266
 DB 864 TGATTCACAAAGTAAACCCCTGTCTGTGTAATCCACTCTTTTCCAGCCCTTCGGCAAA 805
 QY 267 GGTCCATAAGGTCATTTCTGACGCTAACATTAAGCCGAATAGAGGTATCGTTGTTA 326
 DB 804 GGTCCATAAGGTCATTTCTGACGCTAACATTAAGCCGAATAGAGGTATCGTTGTTA 745
 QY 327 AACTGTTGCAATAGCTCCCTGCAATTTGTTTCGCTGAGATGCCACTGNTCAAAATTTCGG 386
 DB 744 AACTGTTGCAATAGCTCCCTGCAATTTGTTTCGCTGAGATGCCACTGNTCAAAATTTCGG 685
 QY 387 ACAGATTCATGAGTGGCCGCTGCCAGCTTTGTTGTTGTTGTTGTTGTTGTTGTTA 446
 DB 684 ACAGATTCATGAGTGGCCGCTGCCAGCTTTGTTGTTGTTGTTGTTGTTGTTGTTA 625
 QY 447 ACTCTCTACGATCATCACTCTTCTTCCCACTGAGGTAAAGTTGTTGTTCTCTCC 506
 DB 624 ACTCTCTACGATCATCACTCTTCTTCCCACTGAGGTAAAGTTGTTGTTCTCTCC 565
 QY 507 CATGTGTTAGTTCGAGCATTAATGACCAATTAAGCCGAATAGAGGTATCGTTGTTA 566
 DB 564 CATGTGTTAGTTCGAGCATTAATGACCAATTAAGCCGAATAGAGGTATCGTTGTTA 505
 QY 567 TGTTCACACCCCTGATGGCAANGNTTCTGTGTTGGCAATTCACAGCTGCTGCCGNACA 624
 DB 504 TGTTCACCA-CCCTGATGGCAANGNTTCTGTGTTGGCAATTCACAGCTGCTGCCGNACA 448

PN US2004191220-A1.
 XX 30-SEP-2004.
 XX 11-JUL-2003; 2003US-00618408.
 XX 11-JUL-2002; 2002US-0395358P.
 XX (EINA/) EINAT P.
 XX (DEIS/) DEISS L.
 XX (MAYA/) MAYA R.
 XX Einat P, Deiss L, Maya R;
 XX WPI; 2004-698655/68.
 XX P-PSDB; ADS00021.
 XX Use of an inhibitor of the WWP1 polypeptide for treating an apoptosis-
 PT related disease in a subject such as cancer, and for preparing a
 PT medicament for treating the apoptosis-related disease.
 XX Disclosure; SEQ ID NO 1; 32pp; English.
 XX The invention relates to a method of treating an apoptosis-related
 CC disease in a subject which comprises administering to the subject a
 CC therapeutic amount of an inhibitor of the WWP1 polypeptide to inhibit
 CC WWP1 so as to thus treat the subject. The method is useful for treating
 CC an apoptosis-related disease, i.e. cancer. Inhibitor of human WWP1
 CC polypeptide is useful in preparing a medicament for treating cancer. The
 CC present sequence represents the human WWP1 gene.
 XX Sequence 2793 BP; 921 A; 523 C; 608 G; 741 T; 0 U; 0 Other;
 SQ
 Query Match 90.7%; Score 566.2; DB 13; Length 2793;
 Best Local Similarity 97.3%; Pred. No. 3.5e-168;
 Matches 582; Conservative 0; Mismatches 15; Indels 1; Gaps 1;
 QY 27 CCCTTTAGTNCAGATGACTTCCCATTCGAGGATCTTTGAATGTTGTTCTTGTTA 86
 Db 1610 CCTTTAGTTACAGATGACTTCCCATTCGAGGATCTTTGAATGTTGTTCTTGTTA 1551
 QY 87 TGATCAACAAGTACCTTACACCTTCACGATATATCTAATTTCCAGCCTTCGCGAGG 146
 Db 1550 TGATCAACAAGTACCTTACACCTTCACGATATATCTAATTTCCAGCCTTCGCGAGG 1491
 QY 147 GGTTCCTTCATCTGTAGGCTTGGATCTTGGATCTTCCACCTGGGTGTTTGTGTTA 206
 Db 1490 GGTTCCTTCATCTGTAGGCTTGGATCTTGGATCTTCCACCTGGGTGTTTGTGTTA 1431
 QY 207 TGATTCACAAGTAAACCCCTGCTGTGTGAATCCACTCTTTTTCACGCTGGTGCAAA 266
 Db 1430 TGATTCACAAGTAAACCCCTGCTGTGTGAATCCACTCTTTTTCACGCTGGTGCAAA 1371
 QY 267 GGTCCATAAGGGTCATTTCTGCGAGCTAAACATTGAAGCCGAATAGAGGTATCGTTGTTA 326
 Db 1370 GGTCCATAAGGGTCATTTCTGCGAGCTAAACATTGAAGCCGAATAGAGGTATCGTTGTTA 1311
 QY 327 AACTGTTGCATAGCTCCCTGCAATTGCTCCGCTGAGATTGCCACTGNTCAAAATTNCGG 386
 Db 1310 AACTGTTGCATAGCTCCCTGCAATTGCTCCGCTGAGATTGCCACTGNTCAAAATTNCGG 1251
 QY 387 ACAGATTCCATGGTAGGCGCTGCGAGCTTGTGTTCTGCTGTTATGATCCACATAATAA 446
 Db 1250 ACAGATTCCATGGTAGGCGCTGCGAGCTTGTGTTCTGCTGTTATGATCCACATAATAA 1191
 QY 447 ACTCCTCTACGATCATCAACTCTTCTTTCACACCTGGAGGTAAAGTTGGTCTCTCC 506
 Db 1190 ACTCCTCTACGATCATCAACTCTTCTTTCACACCTGGAGGTAAAGTTGGTCTCTCC 1131
 QY 507 CATGTGTAGTTCGAGCATTTATGATCCACATAATAGCTCTACCATGAGGATCTTTTCTT 566
 Db 1130 CATGTGTAGTTCGAGCATTTATGATCCACATAATAGGTTCTACCATGAGGATCTTTTCTT 1071

QY 567 TGTTCACACCCCTGATGGCAANGTTCTGTGTGGCATTTCCAGACTGCTGCCGNACA 624
 Db 1070 TGTTCACA-CCCTGATGGCAAGGTTCTGTGTGGCATTTCCAGACTGCTGCCGTACA 1014
 RESULT 9
 AAC90422/c
 ID AAC90422 standard; cDNA; 2827 BP.
 XX AAC90422;
 XX 19-MAR-2001 (first entry)
 XX Human clone 811a coding sequence.
 XX Human; angiogenesis; cancer; Drosophila suppressor of deltex; Su(dx);
 KW CADASIL; wound healing; rheumatoid arthritis; vascular disease;
 KW arteriosclerosis; ss.
 XX Homo sapiens.
 XX WO200073329-A2.
 XX 07-DEC-2000.
 XX 23-MAY-2000; 2000WO-GB001990.
 XX 26-MAY-1999; 99GB-00012132.
 XX (UYMA-) UNIV VICTORIA MANCHESTER.
 XX Baron M;
 XX WPI; 2001-061509/07.
 XX P-PSDB; AAB50048.
 XX Use of homologs of Drosophila Notch regulator gene and encoded protein
 PT products and antibodies in diagnosis and therapy of breast cancer,
 PT angiogenesis and diseases associated with abnormal notch signaling.
 XX Example; Page 28-30; 44pp; English.
 XX The present invention relates to a human homolog of Drosophila suppressor
 CC of deltex (Su(dx)) coding sequence and protein (see AAC90423 and
 CC AAB50049). The human homologs are useful for in vitro diagnosis or
 CC therapy of diseases such as angiogenesis, colon cancer, cervical cancer,
 CC breast cancer, squamous adenocarcinoma, seminoma, melanoma, lung cancer,
 CC dementia, cerebral autosomal dominant arteriopathy with sub-cortical
 CC infarcts and leukoencephalopathy (CADASIL), wound healing, rheumatoid
 CC arthritis, vascular diseases such as arteriosclerosis. The present
 CC sequence is human clone 811a. This sequence has high homology to
 CC Drosophila Su(dx). This sequence was used to screen a human breast tumour
 CC cDNA library, resulting in identification of the human homolog of
 CC AAC90423
 XX
 SQ Sequence 2827 BP; 917 A; 522 C; 578 G; 810 T; 0 U; 0 Other;
 Query Match 90.7%; Score 566.2; DB 4; Length 2827;
 Best Local Similarity 97.3%; Pred. No. 3.5e-168;
 Matches 582; Conservative 0; Mismatches 15; Indels 1; Gaps 1;
 QY 27 CCCTTTAGTNCAGATGACTTCCCATTCGAGGATCTTTGAATGTTGTTCTTGTTA 86
 Db 1043 CCTTTAGTTACAGATGACTTCCCATTCGAGGATCTTTGAATGTTGTTCTTGTTA 984
 QY 87 TGATCAACAAGTACCTTACACCTTCACGATATATCTAATTTCCAGCCTTCGCGAGG 146
 Db 983 TGATCAACAAGTACCTTACACCTTCACGATATATCTAATTTCCAGCCTTCGCGAGG 924
 QY 147 GGTTCCTTCATCTGTAGGCTTCCAGTCTTGGATCTTCCACCTGGGTGTTTGTGTTA 206
 Db 923 GGTTCCTTCATCTGTAGGCTTCCAGTCTTGGATCTTCCACCTGGGTGTTTGTGTTA 864

QY 207 TGATTCAAAAGTAACCCCTGCTGTGTGAATCCACTCTTTTCCAGAGCTGTGGCAA 266
DB 863 TGATTCAAAAGTAACCCCTGCTGTGTGAATCCACTCTTTTCCAGAGCTGTGGCAA 804
QY 267 GGTCCATAAGGGTCATTTCTGCAGCTAATGTAAGCCGAATAGAGGTATCGTTGGTTA 326
DB 803 GGTCCATAAGGGTCATTTCTGCAGCTAATGTAAGCCGAATAGAGGTATCGTTGGTTA 744
QY 327 AACTGTTGATAGTCCCTGCAATTGGTTCGGTGAATGTCACATGNTCAAAATTNCGS 386
DB 743 AACTGTTGATAGTCCCTGCAATTGGTTCGGTGAATGTCACATGNTCAAAATTNCGS 584
QY 387 ACAGATTCCATGTTAGGCGCGCTGCCAGTGTGTGTTCTGTTGTTATGATCCACATAATA 446
DB 683 ACAGATTCCATGTTAGGCGCGCTGCCAGTGTGTGTTCTGTTGTTATGATCCACATAATA 624
QY 447 ACTCCTCTAGGATCATCACTCTTCTTCCAACTCTGGAGGTAAAGGTTGTGTTCTCTCC 506
DB 623 ACTCTTCTAGGATCATCACTCTTCTTCCAACTCTGGAGGTAAAGGTTGTGTTCTCTCC 564
QY 507 CATGTGTTAGTTCGAGCATTTATGANCACACATAATAGCTCTACCATGAGGATCTTTTCTT 566
DB 563 CATGTGTTAGTTCGAGCATTTATGATCCATATAGGTTCTACCATGAGGATCTTTTCTT 504
QY 567 TGTTCACCCCTCTGATGGCAANGTTCGTGTGGCAATCCAGACTGCTGCCGNACA 624
DB 503 TGTTCACCA-CCCTGATGGCAAGTTCGTGTGGCAATCCAGACTGCTGCCGTACA 447

RESULT 10

ABA93745/c
ID ABA93745 standard; cDNA; 3382 BP.

AC ABA93745;

DT 30-APR-2002 (first entry)

XX Human signal transduction cDNA clone tes3_11d21.

DE Human; foetal brain; foetal kidney; melanoma; testis; amygdala;
KW Gene therapy; ss.

XX Homo sapiens.

XX WO200198454-A2.

PN 27-DEC-2001.

PD 25-APR-2001; 2001WO-IB002050.

PF 25-APR-2000; 2000US-0199380P.

XX (GEHU-) GERMAN HUMAN GENOME PROJECT.

XX Wiemann S;

XX WPI; 2002-055860/07.

DR P-PSDB; ABB05708.

XX Human cDNA sequences and clones derived from human fetal brain, fetal
PT kidney, melanoma, testis and amygdala cDNA libraries, useful in genetic
PT screening and therapy.

PS Claim 1; Page 290-291; 611pp; English.

XX The present invention describes assemblages and computer readable media
CC comprising novel human cDNA sequences and clones derived from human
CC foetal brain, foetal kidney, melanoma, testis and amygdala cDNA
CC libraries. ABA93702 to ABA93766 represent human cDNA sequences from the
CC present invention which encode the proteins given in ABB05662 to
CC ABB05729. The human cDNA sequences and clones can be used in gene
CC therapy. The clones may be used in a variety of applications, for example
CC they may be used in profiling assays, for providing large arrays of human

CC genetic material for implementing large-scale screening strategies and
CC for treating diseases via gene therapy procedures
XX
SQ Sequence 3382 BP; 1115 A; 617 C; 705 G; 945 T; 0 U; 0 Other;
Query Match 90.7%; Score 566.2; DB 6; Length 3382;
Best Local Similarity 97.3%; Pred. No. 3.8e-168;
Matches 582; Conservative 0; Mismatches 15; Indels 1; Gaps 1;
QY 27 CCCTTTAGTNCAGATGACCTTCCCATTTCCAGAGATCTTTCAATGTTGTTCTTTGTTA 86
DB 1620 CTTTAGTTACAGATGACTTCCCATTTCCAGAGATCTTTCAATGTTGTTCTTTGTTA 1561
QY 87 TGATCAACAAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGAGCTTCTGGCAGG 146
DB 1560 TGATCAACAAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGAGCTTCTGGCAGG 1501
QY 147 GGTTCCTTCACTGTAAAGCCTTGAGTCTTGGAATCTTCCACATGGTGTGTTGTTGTTA 206
DB 1500 GGTTCCTTCACTGTAAAGCCTTGAGTCTTGGAATCTTCCACATGGTGTGTTGTTGTTA 1441
QY 207 TGATTCACAAAGTAAACCCCTGTCTGTGAATCCACTCTTTTCCAGAGCTGTGGCAA 266
DB 1440 TGATTCACAAAGTAAACCCCTGTCTGTGAATCCACTCTTTTCCAGAGCTGTGGCAA 1381
QY 267 GGTTCATTAAGGTCATTTTCTGCAGCTAACATTTGAAGCCGAATAGAGGTATCGTTGTTA 326
DB 1380 GGTTCATTAAGGTCATTTTCTGCAGCTAACATTTGAAGCCGAATAGAGGTATCGTTGTTA 1321
QY 327 AACTGTTGATAGTCCCTGCAATTGGTTCGGTGAATGTCACATGNTCAAAATTNCGS 386
DB 1320 AACTGTTGATAGTCCCTGCAATTGGTTCGGTGAATGTCACATGNTCAAAATTNCGS 1261
QY 387 ACAGATTCCATGTTAGGCGCGCTGCCAGTGTGTGTTCTGTTGTTATGATCCACATAATA 446
DB 1260 ACAGATTCCATGTTAGGCGCGCTGCCAGTGTGTGTTCTGTTGTTATGATCCACATAATA 1201
QY 447 ACTCCTCTAGGATCATCACTCTTCTTCCCACTCTGGAGGTAAAGGTTGTGTTCTCTCC 506
DB 1200 ACTCCTCTAGGATCATCACTCTTCTTCCCACTCTGGAGGTAAAGGTTGTGTTCTCTCC 1141
QY 507 CATGTGTTAGTTCGAGCATTTATGANCACACATAATAGCTCTACCATGAGGATCTTTTCTT 566
DB 1140 CATGTGTTAGTTCGAGCATTTATGATCCACATAATAGGTTCTACCATGAGGATCTTTTCTT 1081
QY 567 TGTTCACCCCTCTGATGGCAANGTTCGTGTGGCAATCCAGACTGCTGCCGNACA 624
DB 1080 TGTTCACCA-CCCTGATGGCAAGTTCGTGTGGCAATCCAGACTGCTGCCGTACA 1024

RESULT 11

ACN38827/c
ID ACN38827 standard; cDNA; 943 BP.

XX ACN38827;

XX 18-NOV-2004 (first entry)

DE Tumour-associated antigenic target (TAT) cDNA DNA325120, SEQ ID NO:2583.

XX Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
KW chromosome identification; chromosome mapping; gene mapping;
KW gene therapy; cytostatic; gene; ss.

XX Homo sapiens.

XX WO2004030615-A2.

XX 15-APR-2004.

XX	29-SEP-2003; 2003WO-US028547.	QY	457	GATCATCAACTCTCTTTCCACCTGGAGGTAAAGTTGGTCTCTCCCATGTGGTAG	516
XX	02-OCT-2002; 2002US-0414971P.	DB	522	GATCATCAACTCTCTTTCCACCTGGAGGTAAAGTTGGTCTCTCCCATGTGGTAG	463
XX	(GETH) GENENTECH INC.	QY	517	TTGAGCATATTATGANCACATAATACGCTCTACCATGAGGATCTTTTCTTTGTTCCACCC	576
XX	Wu TD, Zhang Z, Zhou Y;	DB	462	TTGAGTATTATGATCCACATATAGTTCTACCATGAGGATCTTTCTTTGTTTCCCA-C	404
XX	WPI; 2004-347921/32.	QY	577	CCTGATGGCAANGNTTCTGTGGCAATCCCGAGATCTGCTGCCGNACA	624
XX	New tumor-associated antigenic target polypeptides and nucleic acids,	DB	403	CCTGATGGCAAGGTTTCTGTGTGGCAATCCCGAGATCTGCTGCCGTACA	356
XX	useful in preparing a medicament for treating or detecting a				
XX	proliferative disorder, e.g. breast, lung, colorectal, ovarian or				
XX	prostate cancer or tumor.				
XX	Claim 1; SEQ ID NO 2583; 7273pp; English.				
XX	The invention relates to human tumour-associated antigenic target (TAT)				
XX	polypeptides, and their related nucleic acids. The TAT polypeptides are				
XX	overexpressed in cancer tissues compared to normal tissues, and may thus				
XX	serve as effective targets for the diagnosis and treatment of cancer in				
XX	mammals. The invention also relates to nucleic acid and polypeptide				
XX	sequences at least 80% identical to the TAT nucleic acids and				
XX	polypeptides; expression vectors and host cells comprising a TAT nucleic				
XX	acid; an antibody specific for a TAT polypeptide; a peptide or organic				
XX	molecule which binds to a TAT polypeptide; fusion proteins comprising a				
XX	TAT polypeptide; and methods and compositions for the treatment or				
XX	diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,				
XX	antibodies, antagonists, binding molecules and compositions are useful				
XX	for diagnosing or treating a cell proliferative disorder associated with				
XX	increased TAT expression, particularly cancers such as breast cancer,				
XX	colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder				
XX	cancer, pancreatic cancer, cervical cancer, cancers of the central				
XX	nervous system, melanoma and leukaemia. TAT nucleic acids may further be				
XX	used as hybridisation probes, in chromosome and gene mapping, in				
XX	chromosome identification and in gene therapy. The present sequence				
XX	represents a TAT nucleic acid of the invention				
XX	Sequence 943 BP; 309 A; 195 C; 205 G; 234 T; 0 U; 0 Other;				
XX	Query Match 90.3%; Score 563.6; DB 13; Length 943;				
XX	Best Local Similarity 98.1%; Pred. No. 1.4e-167;				
XX	Matches 577; Conservative 0; Mismatches 10; Indels 1; Gaps 1;				
QY	37 CAGATGACTTCCCATTCGAGGATCTTTGAATGTTGTTCTGTGTTATGATCAACA	96			
DB	942 CAGATGACTTCCCATTCGAGGATCTTTGAATGTTGTTCTGTGTTATGATCAACA	883			
QY	97 AGTACCTTACACCTTCACGAGTATATCTAATTTCCAGCCTTCGCGAGGGTTCTTCAT	156			
DB	882 AGTACCTTACACCTTCACGAGTATATCTAATTTCCAGCCTTCGCGAGGGTTCTTCAT	823			
QY	157 TCTGTAAGCCTTGAGTCTTGGATCTTCCCATCGGTTGTTTGTGTTATGATCAACA	216			
DB	822 TCTGTAAGCCTTGAGTCTTGGATCTTCCCATCGGTTGTTTGTGTTATGATCAACA	763			
QY	217 AGTAAACCTGCTGTTGTAATCCACTCTTTTCCAGCCTTCGCGAGGGTTCCATAAG	276			
DB	762 AGTAAACCTGCTGTTGTAATCCACTCTTTTCCAGCCTTCGCGAGGGTTCCATAAG	703			
QY	277 GGTCAATTTCTGACGCTAAACATTAAGCCGATAGAGGTATCGTTGGTTAACTGTTGCA	336			
DB	702 GGTCAATTTCTGACGCTAAACATTAAGCCGATAGAGGTATCGTTGGTTAACTGTTGCA	643			
QY	337 TAGTCCCTGCAATGGTTCGCTGAGATGCCACTGNTCAAATTCGGACAGATCCA	396			
DB	642 TAGTCCCTGCAATGGTTCGCTGAGATGCCACTGNTCAAATTCGGACAGATCCA	583			
QY	397 TGGTAGCCGCTGCCAGCTGTTGTTCTGCGTTTATGATCCACATATTAACCTCTAC	456			
DB	582 TGGTAGCCGCTGCCAGCTGTTGTTCTGCGTTTATGATCCACATATTAACCTCTCTAC	523			
XX	RESULT 12				
XX	AAC90423/c				
XX	ID AAC90423 standard; cDNA; 3495 BP.				
XX	AC AAC90423;				
XX	XX 19-MAR-2001 (first entry)				
XX	DE Human homolog of Drosophila suppressor of deltex coding sequence.				
XX	KW Human; angiogenesis; cancer; Drosophila suppressor of deltex; Su(dx);				
XX	KW CADASIL; wound healing; rheumatoid arthritis; vascular disease;				
XX	OS arteriosclerosis; ss.				
XX	XX Homo sapiens.				
XX	PN W0200073329-A2.				
XX	XX 07-DEC-2000.				
XX	XX 23-MAY-2000; 2000WO-GB001990.				
XX	XX 26-MAY-1999; 99GB-00012132.				
XX	XX (UYMA-) UNIV VICTORIA MANCHESTER.				
XX	XX Baron M;				
XX	XX WPI; 2001-061509/07.				
XX	XX P-PSDB; AAB50049.				
XX	XX Use of homologs of Drosophila Notch regulator gene and encoded protein				
XX	XX products and antibodies in diagnosis and therapy of breast cancer,				
XX	XX angiogenesis and diseases associated with abnormal notch signaling.				
XX	XX Claim 2; Page 30-33; 44pp; English.				
XX	XX The present invention relates to a human homolog of Drosophila suppressor				
XX	XX of deltex (Su(dx)) coding sequence and protein (AAC90423 and AAB50049).				
XX	XX The human homologs are useful for in vitro diagnosis or therapy of				
XX	XX diseases such as angiogenesis, colon cancer, cervical cancer, breast				
XX	XX cancer, squamous adenocarcinoma, seminoma, melanoma, lung cancer,				
XX	XX dementia, cerebral autosomal dominant arteriopathy with sub-cortical				
XX	XX infarcts and leukoencephalopathy (CADASIL), wound healing, rheumatoid				
XX	XX arthritis, vascular diseases such as arteriosclerosis				
XX	XX Sequence 3495 BP; 1155 A; 636 C; 729 G; 975 T; 0 U; 0 Other;				
XX	XX Query Match 88.5%; Score 552.2; DB 4; Length 3495;				
XX	XX Best Local Similarity 97.0%; Pred. No. 1.1e-163;				
XX	XX Matches 580; Conservative 0; Mismatches 15; Indels 3; Gaps 2;				
QY	27 CCCTTAGTNCAGATGACTTCCCATTCGAGGATCTTTGAATGTTGTTCTGTGTTA	86			
DB	1710 CCCTTAGTNCAGATGACTTCCCATTCGAGGATCTTTGAATGTTGTTCTGTGTTA	1653			
QY	87 TGATCAACAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGCCTTCCTGGCAGG	146			
DB	1652 TGATCAACAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGCCTTCCTGGCAGG	1593			

147 GGTCTTCAATCTGTAGCCCTTGAGTCTTGGATCTCCACCTGGGTTGTTTGTGTTA 206
 1592 GGTCTTCAATCTGTAGCCCTTGAGTCTTGGATCTCCACCTGGGTTGTTTGTGTTA 1533
 207 TGATTCACAAAGTAAACCCCTGTCTGTGTAATCCACTCTTTTCCAGACCTGTGTGCAAA 266
 1532 TGATTCACAAAGTAAACCCCTGTCTGTGTAATCCACTCTTTTCCAGACCTGTGTGCAAA 1473
 267 GGTCCATAAGGGTCATTTCTGCAGCTAACATTTGAGCCGCAATAGAGGTATCGTTGGTTA 326
 1472 GGTCCATAAGGGTCATTTCTGCAGCTAACATTTGAGCCGCAATAGAGGTATCGTTGGTTA 1413
 327 AACTGTTGTCATAGCTCCCTGCAATTTGGTTCGGCTGAGATTGCCACTGNTCAAAATTNCGG 386
 1412 AACTGTTGTCATAGCTCCCTGCAATTTGGTTCGGCTGAGATTGCCACTGTTCAAAATTNCGG 1353
 387 ACAGATTCCATGTTAGCCCTGCGCAGCTGTTGTTGTTCTGTTGTTATGATCCACATAATA 446
 1352 ACAGATTCCATGTTAGCCCTGCGCAGCTGTTGTTGTTCTGTTGTTATGATCCACATAATA 1293
 447 ACTCCTCTACGATCATCAACTCTTTCTTCCCACTGGAGTAAAGGTTGTTGTTCTCTCC 506
 1292 ACTCCTCTACGATCATCAACTCTTTCTTCCCACTGGAGTAAAGGTTGTTGTTCTCTCC 1233
 507 CATGTGTGATGTTGAGCAATTTATGATCCACATAATAAGCTCTACCATGAGGATCTTTTCTT 566
 1232 CATGTGTGATGTTGAGCAATTTATGATCCACATAATAAGCTCTACCATGAGGATCTTTTCTT 1173
 567 TGTTCACACCCCTGATGGCAAGNNTCTGTGTTGGCAATCCAGACTGCTGCCGNACA 624
 1172 TGTTCACCA-CCCTGATGGCAAGGTTCTGTGTTGGCAATCCAGACTGCTGCCGNACA 1116

RESULT 13

ADP84403/c

ID ADP84403 standard; DNA; 2305 BP.

XX AC ADP84403;

XX 09-SEP-2004 (first entry)

XX Human breast-specific protein coding sequence #2.

XX human; breast-specific protein; breast cancer; gene; ds.

XX Homo sapiens.

XX WO2004053077-A2.

XX 24-JUN-2004.

XX 05-DEC-2003; 2003WO-US038815.

XX 05-DEC-2002; 2002US-0431123P.

XX (DIAD-) DIADEXUS INC.

XX Macina RA, Turner LR, Sun Y, Chen H, Rodriguez M;

XX WPI; 2004-468848/44.

XX P-PSDB; ADP84498, ADP84499.

XX New breast specific nucleic acid molecules and polypeptides useful for
 PT diagnosing, preventing or treating breast cancer, for producing
 PT transgenic animals or cells, or for research purposes.

XX Claim 1; SEQ ID NO 2; 521pp; English.

XX The invention comprises the amino acid and coding sequences of human
 CC breast-specific proteins. The DNA and protein sequences of the invention
 CC are useful for the diagnosis, treatment and prevention of breast cancer.
 CC the present DNA sequence encodes a human breast-specific protein of the
 CC invention.

XX SQ Sequence 2305 BP; 712 A; 571 C; 545 G; 476 T; 0 U; 1 Other;
 Query Match 73.2%; Score 456.8; DB 12; Length 2305;
 Best Local Similarity 96.0%; Pred. No. 1.4e-133;
 Matches 486; Conservative 0; Mismatches 18; Indels 2; Gaps 2;
 QY 120 TATCTAATTTCCAGCCTTCTGSC-AGGGGTTCTTCAATCTGTAAAGCCTTGAGTCTTGG 178
 Db 1868 TCTAAATTGTGCGAGCCTTCTGCCAAGGGTTTCTTCAATCTGTAAAGCCTTGAGTCTTGG 1809
 QY 179 ATCTTCCCACTGGGTTGTTTGTGTTATGATTCACAAAAGTAAACCCCTGTCTGTGGAATC 238
 Db 1808 ATCTTCCCACTGGGTTGTTTGTGTTATGATTCACAAAAGTAAACCCCTGTCTGTGGAATC 1749
 QY 239 CACTCTTTTTCAGCCCTGTGGCAAGGTCCATAAGGGTCAATTTCTGACGCTAACAT 298
 Db 1748 CACTCTTTTTCAGCCCTGTGGCAAGGTCCATAAGGGTCAATTTCTGACGCTAACAT 1689
 QY 299 TGAAGCCGAATAGAGGTATCGTTGGTTAAACTGTTGCATAGCTCCCTGCAATTTGGTTCCG 358
 Db 1688 TGAAGCCGAATAGAGGTATCGTTGGTTAAACTGTTGCATAGCTCCCTGCAATTTGGTTCCG 1629
 QY 359 CTGAGATTGCCACTGNTCAAAATTNCGGACAGATTCATGGTAGGCCCTGCCACGTTGT 418
 Db 1628 CTGAGATTGCCACTGTTCAAAATTTTCGGACAGATTCATGGTAGGCCCTGCCACGTTGT 1569
 QY 419 TGTCTGTGTTTATGATCCACATAATAAACTCCTCTACGATCATCAACTCTTCTTTCCCA 478
 Db 1568 TGTCTGTGTTTATGATCCACATAATAAACTCCTCTACGATCATCAACTCTTCTTTCCCA 1509
 QY 479 ACTCGAGGTAAAGGTTGTGTTCTCTCCCATGTTGGTAGTTCGAGCATTTATGACCACATA 538
 Db 1508 ACTCGAGGTAAAGGTTGTGTTCTCTCCCATGTTGGTAGTTCGAGCATTTATGACCACATA 1449
 QY 539 ATACGCTCTACCATGAGGATCTTTTCTTTGTTTCCCACTGAGTGAAGCAANGNTTCTGTGT 598
 Db 1448 ATAGGTTCTACCATGAGGATCTTTTCTTTGTTTCCCA-CCCTGATGGCAAGGTTCTGTGT 1390
 QY 599 TGGCATTTCCAGACTGCTGCCGNACA 624
 Db 1389 TGGCATTTCCAGACTGCTGCCGNACA 1364

RESULT 14

AAS80664/c

ID AAS80664 standard; cDNA; 2217 BP.

XX AC AAS80664;

XX 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #16468.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR P-PSDB; ABG16477.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 16468; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pst_sequences
XX
SQ Sequence 2217 BP; 708 A; 434 C; 505 G; 570 T; 0 U; 0 Other;
Query Match 39.5%; Score 246.6; DB 5; Length 2217;
Best Local Similarity 64.3%; Pred. No. 5.2e-67;
Matches 388; Conservative 0; Mismatches 200; Indels 15; Gaps 1;
QY 12 TAAGCATTGTGGACCCCTTTAGTNCAGATGACCTCCCATGCGAGGATCTTTGAATGTT 71
DB 1079 TAGGCTATCTGAGGTCCATGTTCTAGGCGAGATTTCTCTGCGGGATCTATATAGTG 1020
QY 72 GTTGTCTTCTGTGTATGATCAACAAAGTACCTTACACCTTCCAGAGTATATCTAATTTCC 131
DB 1019 GTAGTCTTCTTATTTGTGGTCCCAAAATATGGAATTCATCCCACTGTGAATCTCATTTCC 960
QY 132 CAGCCTTCTGGCAGGGGTCTTCTCATTTCTGTGAAGCTTGGATCTTCCACTGG 191
DB 959 CAACCTTCAGTAAGGCTTTTCATTTAATGACCTTGACTTCTGGGGTCTTCCATTGT 900
QY 192 GTTGTCTTCTGTGTATGATCAACAAAGTAAACCTGCTGTGTGAATCCACTCTTTTTC 251
DB 899 GTAAATTCGTGTGTGTGTTGACGAAATATACCTGCAATGCTGTCTGTCTCTCTCC 840
QY 252 CAGCCTGCTGCGAAGTCCATGAAGGTCTATTTCT-----GCAGCTAAC 296
DB 839 CATCCAGGTGCAATGACCAAGAGGATCAAAATTTCTTACTTGTGTGAATGAGCAATAA 780
QY 297 ATTGAAGCCCAATAGAGGTATCGTTGTGTTAACTGTTGTCATAGCTCCCTGCAATGGTTC 356
DB 779 TCTGTATCCCATTAATGAATCTCTGTTTAACTGCTGCAATCTCTCTGAAGCTGACTA 720
QY 357 CGCTGAGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 416
DB 719 CGCTGTAGTGCCTATGTTTCATAGTTCGCGAGGATTCAGTGTGTCCTGCGAGTT 660
QY 417 GTTGTCTTCTGTGTATGATCAACAAATTAATCACTCTCTACGATCATCACTCTTCTTCC 476
DB 659 GTTGTCTTCTGTAATGCTCAACATAATAATACGCTCCCATGTTGTGTCAACCCGCGTTC 600
QY 477 CAACCTGAGGTAAAGGTGTGGTCTCTCCCATGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 536
DB 599 CAGCCAGGAGGTAGAGGTCTTGGTCTATCCCAATGTTCTTCTTCTCAACATGATCTACA 540

QY 537 TAATAGCTCTACGATGAGGATCTTTCTTTTCTCCACCCCTGATGCGAANGTTCTGT 596
DB 539 TAGTAACTCGCCGCTGCTGCTCCACTCTCTGCTCCCAACGAGTGGCAAGGAGCTTGA 480
QY 597 GTT 599
DB 479 GTT 477

RESULT 15
ABA82679/c
ID ABA82679 standard; DNA; 2377 BP.
XX ABA82679;
AC ABA82679;
XX
DT 25-JAN-2002 (first entry)
DE Atrophin-1 interacting protein (AIP4) gene SEQ ID NO:65.
XX Human; high bone mass; HBM gene; Zmax1 gene; chromosome 11; 11q13.3;
XX sequence tagged site; STS; osteoporosis; osteopathic; gene therapy;
XX antisense therapy; vaccine; bone disorder; Paget's disease; sclerostosis;
XX osteomalacia; fibrous dysplasia; ds.
XX Homo sapiens.
XX WO200177327-A1.
PN 18-OCT-2001.
XX
XX 21-JUN-2000; 2000WO-US016951.
XX
PR 05-APR-2000; 2000US-00543771.
PR 05-APR-2000; 2000US-00544398.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Carulli JP, Little RD, Recker RR, Johnson ML;
XX WPI; 2001-657171/75.
XX
XX New high bone mass (HBM) and Zmax1 genes and proteins useful for
XX modulating bone mass for the treatment of e.g. osteoporosis.
XX
XX Claim 79; Page 367; 443pp; English.
XX
XX The present invention describes the human Zmax1 gene and the high bone
XX mass (HBM) gene, which are found on chromosome 11q13.3. The Zmax1 and HBM
XX genes have osteopathic activities. The genes can be used in gene therapy,
XX antisense therapy and in the production of vaccines. They can be used in
XX the diagnosis and treatment of bone disorders including osteoporosis,
XX Paget's disease, sclerostosis, osteomalacia and fibrous dysplasia.
XX ABA82038 to ABA82700 and AAG68168 to AAG68193 represent sequences used in
XX the exemplification of the present invention
XX
SQ Sequence 2377 BP; 765 A; 458 C; 532 G; 622 T; 0 U; 0 Other;
Query Match 39.5%; Score 246.6; DB 5; Length 2377;
Best Local Similarity 64.3%; Pred. No. 5.4e-67;
Matches 388; Conservative 0; Mismatches 200; Indels 15; Gaps 1;
QY 12 TAAGCATTGTGGACCCCTTTAGTNCAGATGACCTCCCATGCGAGGATCTTTGAATGTT 71
DB 1079 TAGGCTATCTGAGGTCCATGTTCTAGGCGAGATTTCTCTGCGGGATCTATATAGTG 1020
QY 72 GTTGTCTTCTGTGTATGATCAACAAAGTACCTTACACCTTCCAGAGTATATCTAATTTCC 131
DB 1019 GTAGTCTTCTTATTTGTGGTCCCAAAATATGGAATTCATCCCACTGTGAATCTCATTTCC 960
QY 132 CAGCCTTCTGGCAGGGGTCTTCTCATTTCTGTGAAGCTTGGATCTTCCACTGG 191
DB 959 CAACCTTCAGTAAGGCTTTTCATTTAATGACCTTGACTTCTGGGGTCTTCCATTGT 900

QY	192	GTGTTTGTGTTATGATTCAAAAGTAAACCTGTCTGTGTGAATCCACTCTTTTTC	251
Db	899	GTAAATCGTGTGTGTGGTTGACGAAATATACTCTGCCATTGCTGTCTTCTTCTCC	840
QY	252	CAGCCTGGTGGCAAGGTCCATAAGGGTCATTTTCT-----GCAGCTAAC	296
Db	839	CATCCAGGTGGCAATGGACCAAGAGGATCAAATTCCTTACTTTGTGATGTAGCAATAAA	780
QY	297	ATTGAAGCCGAATAGAGGTATCGTTGGTTAAACTGTGTCATAGCTCCCTGCAATTGGTTC	356
Db	779	TCTTGATTCCTATAAATGAATCTCTGGTTAACTGCTGCATTTGCTTGAAGCTGACTA	720
QY	357	CGCTGAGATTGCCACTGNTCAAAATTCGGAACAGATTCATGTAGGCGCTGCCAGGTT	416
Db	719	CGCTGAGTGCATTTGTCATAGTTCGGAGGATTCAGTGTGGCTCTGCCACGTT	660
QY	417	GTGTTCTGTGTATGATCCACATATAAACTCTCTACGATCATCAACTCTTCTTTC	476
Db	659	GTGTTCTGTGAAATGGTCAACATAATAAATACGTCCCATGTTGTCAACCCCGCTTC	600
QY	477	CAACTGGAGGTAAAGGTTGTGGTCTCTCCCATGTGTAGTTCGAGCATTTATGACCACA	536
Db	599	CAGCCAGGAGGTAGAGGTTCTGGTCTATCCCATGTGTCTTTTCTCAACATGATCTACA	540
QY	537	TAATACGCTCTACCATGAGGATCTTTTCTTTGTTCCCAACCCCTGATGGCAANGNTCTGT	596
Db	539	TAGTAAACTCGCCGTGTGTGTCCTCACTCTCTGCTCCCAACCAAGGTGGCAAGGAGCTTGA	480
QY	597	GTT 599	
Db	479	GTT 477	

Search completed: June 30, 2005, 05:38:24
Job time : 478 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2005, 05:25:19 ; Search time 531 Seconds
(without alignments)
7345.241 Million cell updates/sec

Title: US-10-618-408-3
Perfect score: 624
Sequence: 1 cgcgcgctntaagcattt.....tccacagctgctgccnaca 624

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 6067389 seqs, 3125258755 residues

Total number of hits satisfying chosen parameters: 12134778

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:
22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	616	98.7	624	19	US-10-618-408-3
2	566.2	90.7	2052	14	US-10-097-534-3
3	566.2	90.7	2052	15	US-10-185-050-45
4	566.2	90.7	2052	15	US-10-101-510-56
5	566.2	90.7	2052	18	US-10-363-828-3
6	566.2	90.7	2731	17	US-10-188-186-113
7	566.2	90.7	2793	19	US-10-618-408-1

Sequence 2, Appli	US-10-097-534-2	14	2377	39.5	246.6	8
Sequence 65, Appl	US-10-374-979-65	17	2377	39.5	246.6	9
Sequence 383, App	US-10-240-425-383	18	2377	39.5	246.6	10
Sequence 65, Appl	US-10-182-936A-65	18	2377	39.5	246.6	11
Sequence 128, App	US-10-283-975A-128	19	2377	39.5	246.6	12
Sequence 644, App	US-10-477-238A-644	19	2377	39.5	246.6	13
Sequence 644, App	US-10-680-287A-644	20	2377	39.5	246.6	14
Sequence 644, App	US-10-477-173-644	21	2377	39.5	246.6	15
Sequence 234, App	US-09-919-039-234	21	2678	39.5	246.6	16
Sequence 40, Appl	US-10-489-740-40	21	3745	39.5	246.6	17
Sequence 1706, Ap	US-10-956-157-1706	21	3455	32.6	203.6	18
Sequence 4, Appli	US-10-097-534-4	21	3475	32.6	203.6	19
Sequence 10, Appl	US-10-363-828-10	21	3475	32.6	203.6	20
Sequence 2166, Ap	US-10-723-860-2166	20	3475	32.6	203.6	21
Sequence 47, Appl	US-10-185-050-47	20	3476	32.6	203.6	22
Sequence 2090, Ap	US-10-062-674-2090	14	4658	29.0	181	23
Sequence 6195, Ap	US-10-723-860-6195	20	4573	27.1	168.8	24
Sequence 6414, Ap	US-10-723-860-6414	20	4573	27.1	168.8	25
Sequence 100, App	US-09-774-639-100	20	2351	17.9	111.8	26
Sequence 25, Appl	US-09-969-730-25	10	2351	17.9	111.8	27
Sequence 322, App	US-09-764-853-322	17	2351	17.9	111.8	28
Sequence 7, Appli	US-09-764-853-7	9	556	8.8	55	29
Sequence 8, Appli	US-10-097-534-7	14	2844	8.7	54	30
Sequence 203, App	US-10-097-534-8	14	5749	8.7	54	31
Sequence 132, App	US-10-171-581-203	15	5749	8.7	54	32
Sequence 30, Appl	US-10-473-974-132	20	5749	8.7	54	33
Sequence 28, Appl	US-10-149-827A-30	19	711	8.6	53.8	34
Sequence 91, Appl	US-10-149-827A-28	19	1225	8.6	53.8	35
Sequence 29, Appl	US-09-764-853-91	9	2211	8.6	53.8	36
Sequence 1, Appli	US-10-149-827A-29	19	2219	8.6	53.8	37
Sequence 1640, Ap	US-10-058-518-1	16	2256	8.6	53.8	38
Sequence 1640, Ap	US-10-172-118-1640	17	2264	8.6	53.8	39
Sequence 1640, Ap	US-10-342-887-1640	18	2264	8.6	53.8	40
Sequence 3, Appli	US-10-313-955-3	16	2848	8.6	53.4	41
Sequence 1799, Ap	US-10-956-157-1799	21	4929	8.6	53.4	42
Sequence 6296, Ap	US-10-032-585-6296	21	2499	8.4	52.2	43
Sequence 98, Appl	US-10-785-813-98	16	2662	8.3	51.8	44
Sequence 5, Appli	US-10-313-955-5	16	3226	8.0	50	45

ALIGNMENTS

RESULT 1
US-10-618-408-3
; Sequence 3, Application US/10618408
; Publication No. US20040191220A1
; GENERAL INFORMATION:
; APPLICANT: Paz Einat, Louis Deiss, and Ruth Maya
; TITLE OF INVENTION: WMP1 AND USES THEREOF
; FILE REFERENCE: 67722-A; 074/US1
; CURRENT APPLICATION NUMBER: US/10/618,408
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: 60/395358
; PRIOR FILING DATE: 2002-07-11
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 624
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (11)..(11)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36)..(36)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (374)..(374)
; OTHER INFORMATION: n is a, c, g, or t

Sequence 45, Application US/10185050
Publication No. US20030077577A1
GENERAL INFORMATION:
APPLICANT: Pirozzi, Gregorio
Kay, Brian K.
Fowlkes, Dana M.
TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 233
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10185,050
FILING DATE: 28-Jun-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/826,516
FILING DATE: 03-Apr-1997
ATTORNEY/AGENT INFORMATION:
NAME: MISROCK, S. LESLIE
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-208-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 896-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 2052 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-10-185-050-45
Query Match 90.7%; Score 566.2; DB 14; Length 2052;
Best Local Similarity 97.3%; Pred. No. 7.6e-166;
Matches 582; Conservative 0; Mismatches 15; Indels 1; Gaps 1;
QY 27 CCCTTTAGTNCAGATGACTTCCCATTCGCGAGGATCTTTGAATGTTGTTGTTGTTGTTA 86
DB 1044 CCTTGTAGTACAGATGACTTCCCATTCGCGAGGATCTTTGAATGTTGTTGTTGTTA 985
QY 87 TGATCAACAAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGCCCTTCGCGAGG 146
DB 984 TGATCAACAAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGCCCTTCGCGAGG 925
QY 147 GGTTCCTTCATCTGTAAGCCTTCAGTCTTGGATCTTCCACATGGGTTGTTTGTGTTA 206
DB 924 GGTTCCTTCATCTGTAAGCCTTCAGTCTTGGATCTTCCACATGGGTTGTTTGTGTTA 865
QY 207 TGATTCACAAAGTAAACCTCTGTTGTAATCCACTCTTTTTCCAGCCCTTCGTTGCAAA 266
DB 864 TGATTCACAAAGTAAACCTCTGTTGTAATCCACTCTTTTTCCAGCCCTTCGTTGCAAA 805
QY 267 GGTCCATTAAGGGTCATTTCTGCGAGTAACTTGAAGCCGAATAGAGGATATCGTTGTTA 326
DB 804 GGTCCATTAAGGGTCATTTCTGCGAGTAACTTGAAGCCGAATAGAGGATATCGTTGTTA 745
QY 327 AACTGTTGCATAGCTCCCTGCAATTCGTTCCGCTGAGATTCGCACCTGNTCAAAATTCGG 386
DB 744 AACTGTTGCATAGCTCCCTGCAATTCGTTCCGCTGAGATTCGCACCTGNTCAAAATTCGG 685
QY 387 ACAGATTCATGTTAGGCGCGCTGCCACGTTGTTGTTCTGGTGTATGATCCACATAATAA 446
DB 684 ACAGATTCATGTTAGGCGCGCTGCCACGTTGTTGTTCTGGTGTATGATCCACATAATAA 625
QY 447 ACTCTCTACGATCATCACTCTTCTTCCCAACCTGGAGGTAAGGTTGTTGTTCTCTCC 506
DB 624 ACTCTCTACGATCATCACTCTTCTTCCCAACCTGGAGGTAAGGTTGTTGTTCTCTCC 565
QY 507 CATGTGTAGTTCGAGCATTTATGANCACACATATACCTCTACCATGAGGATCTTTTCTT 566
DB 564 CATGTGTAGTTCGAGCATTTATGATCCACATATAGGTTCTACCATGAGGATCTTTTCTT 505
QY 567 TGTTCACACCCCTGATGGCAANGTTCTGTGTGGCAATTCCTCCAGACTGCTGCCGNACA 624
DB 504 TGTTCACCA-CCCTGATGGCAAGGTTCTGTGTGGCAATTCCTCCAGACTGCTGCCGNACA 448

RESULT 4

US-10-101-510-56/c
Sequence 56, Application US/10101510
Publication No. US20030148295A1
GENERAL INFORMATION:
APPLICANT: WAN, JACKSON
APPLICANT: WANG, YIXIN
TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
FILE REFERENCE: 15117.0012
CURRENT APPLICATION NUMBER: US/10/101,510
CURRENT FILING DATE: 2002-03-20
PRIOR APPLICATION NUMBER: 60/276,947
PRIOR FILING DATE: 2001-03-20
NUMBER OF SEQ ID NOS: 805
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 56
LENGTH: 2052
TYPE: DNA
ORGANISM: Homo sapiens
US-10-101-510-56

Query Match 90.7%; Score 566.2; DB 15; Length 2052;
Best Local Similarity 97.3%; Pred. No. 7.6e-166;
Matches 582; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

QY 27 CCCTTTAGTNCAGATGACTTCCCATTCGCGAGGATCTTTGAATGTTGTTGTTGTTA 86
DB 1044 CCTTGTAGTACAGATGACTTCCCATTCGCGAGGATCTTTGAATGTTGTTGTTGTTA 985
QY 87 TGATCAACAAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGCCCTTCGCGAGG 146
DB 984 TGATCAACAAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGCCCTTCGCGAGG 925
QY 147 GGTTCCTTCATCTGTAAGCCTTCAGTCTTGGATCTTCCACATGGGTTGTTTGTGTTA 206
DB 924 GGTTCCTTCATCTGTAAGCCTTCAGTCTTGGATCTTCCACATGGGTTGTTTGTGTTA 865
QY 207 TGATTCACAAAGTAAACCTCTGTTGTAATCCACTCTTTTTCCAGCCCTTCGTTGCAAA 266
DB 864 TGATTCACAAAGTAAACCTCTGTTGTAATCCACTCTTTTTCCAGCCCTTCGTTGCAAA 805
QY 267 GGTCCATTAAGGGTCATTTCTGCGAGTAACTTGAAGCCGAATAGAGGATATCGTTGTTA 326
DB 804 GGTCCATTAAGGGTCATTTCTGCGAGTAACTTGAAGCCGAATAGAGGATATCGTTGTTA 745
QY 327 AACTGTTGCATAGCTCCCTGCAATTCGTTCCGCTGAGATTCGCACCTGNTCAAAATTCGG 386
DB 744 AACTGTTGCATAGCTCCCTGCAATTCGTTCCGCTGAGATTCGCACCTGNTCAAAATTCGG 685
QY 387 ACAGATTCATGTTAGGCGCGCTGCCACGTTGTTGTTCTGGTGTATGATCCACATAATAA 446
DB 684 ACAGATTCATGTTAGGCGCGCTGCCACGTTGTTGTTCTGGTGTATGATCCACATAATAA 625
QY 447 ACTCTCTACGATCATCACTCTTCTTCCCAACCTGGAGGTAAGGTTGTTGTTCTCTCC 506
DB 624 ACTCTCTACGATCATCACTCTTCTTCCCAACCTGGAGGTAAGGTTGTTGTTCTCTCC 565

Db 624 ACTCTTCTACGATCATCAACTCTTCTTCCCAACCTGGAGGTAAGGTTGTGGTCTCTCC 565
QY 507 CATGTGGTAGTTCGACATTATGANCACATAATAGCTCTACCATGAGGATCTTTTCTT 566
Db 564 CATGTGGTAGTTCGAGTATTATGATCCACATAATAGGTTCTACCATGAGGATCTTTTCTT 505
QY 567 TGTTCCTCCACCCCTGATGGCAANGNTTCTGTGTGGCAITCCAGACTGCTGCCGNACA 624
Db 504 TGTTCCTCA-CCCTGATGGCAAGGTTCTGTGTGGCAITCCAGACTGCTGCCGNACA 448
RESULT 5
US-10-363-828-3/c
; Sequence 3, Application US/10363828
; Publication No. US20040076973A1
; GENERAL INFORMATION:
; APPLICANT: Isis Pharmaceuticals, Inc.
; APPLICANT: Brett P. Monia
; APPLICANT: Lex M. Cowbert
; TITLE OF INVENTION: ANTISENSE MODULATION OF UBIQUITIN PROTEIN LIGASE EXPRESSION
; FILE REFERENCE: RTSP-0164
; CURRENT APPLICATION NUMBER: US/10/363,828
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 09/657,481
; PRIOR FILING DATE: 2000-09-07
; NUMBER OF SEQ ID NOS: 93
; SEQ ID NO 3
; LENGTH: 2052
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2052)
US-10-363-828-3
Query Match 90.7%; Score 566.2; DB 18; Length 2052;
Best Local Similarity 97.3%; Pred. No. 7.6e-166;
Matches 582; Conservative 0; Mismatches 15; Indels 1; Gaps 1;
QY 27 CCTTTAGTNCAGATGACTTCCCATTCGCGAGGATCTTTGAATGTGTTCTTCTTGTTA 86
Db 1044 CCTTTAGTACAGATGACTTCCCATTCGCGAGGATCTTTGAATGTGTTCTTCTTGTTA 985
QY 87 TGATCAACAAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGGCTTCTGGCAGG 146
Db 984 TGATCAACAAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGGCTTCTGGCAGG 925
QY 147 GGTTCCTTCAATCTGTAAAGCTTGAGTTCCTGGATCTTCCACCTGGGTTGTTTGTGTTA 206
Db 924 GGTTCCTTCAATCTGTAAAGCTTGAGTTCCTGGATCTTCCACCTGGGTTGTTTGTGTTA 865
QY 207 TGATTCACAAAGTAAACCTGCTGTGTGAATCCACTCTTTTCCAGGCTGTTGGCAAA 266
Db 864 TGATTCACAAAGTAAACCTGCTGTGTGAATCCACTCTTTTCCAGGCTGTTGGCAAA 805
QY 267 GGTTCATTAAGGTCATTTCTGAGCTAACATTTGAAGCCGAATAGAGGATCGTTGTTA 326
Db 804 GGTTCATTAAGGTCATTTCTGAGCTAACATTTGAAGCCGAATAGAGGATCGTTGTTA 745
QY 327 AACTGTGTGATAGCTCCCTGCAATTTGGTTCGGTGAATTTGCCACTGNTCAAAATTCGG 386
Db 744 AACTGTGTGATAGCTCCCTGCAATTTGGTTCGGTGAATTTGCCACTGNTCAAAATTCGG 685
QY 387 ACAGATTCATAGTGTAGGCGCTGCCAGTTGTGTGTTCTGTGTATATGCCACATAATAA 446
Db 684 ACAGATTCATAGTGTAGGCGCTGCCAGTTGTGTGTTCTGTGTATATGCCACATAATAA 625
QY 447 ACTCTCTACGATCATCAACTCTTCTTCCCAACCTGGAGGTAAGGTTGTGTTCTCTCC 506
Db 624 ACTCTCTACGATCATCAACTCTTCTTCCCAACCTGGAGGTAAGGTTGTGTTCTCTCC 565
QY 507 CATGTGGTAGTTCGAGGATATGANCACATAATAGCTCTACCATGAGGATCTTTTCTT 566

Db 564 CATGTGGTAGTTCGAGTATATGATCCACATAATAGGTTCTACCATGAGGATCTTTTCTT 505
QY 567 TGTTCCTCCACCCCTGATGGCAANGNTTCTGTGTGGCAITCCAGACTGCTGCCGNACA 624
Db 504 TGTTCCTCA-CCCTGATGGCAAGGTTCTGTGTGGCAITCCAGACTGCTGCCGNACA 448
RESULT 6
US-10-188-186-113/c
; Sequence 113, Application US/10188186
; Publication No. US20040029789A1
; GENERAL INFORMATION:
; APPLICANT: Anderson et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-397C
; CURRENT APPLICATION NUMBER: US/10/188,186
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: 60/303046
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/360814
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/303828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/323380
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/361133
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/304016
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/304502
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 60/305262
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/373881
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/305673
; PRIOR FILING DATE: 2001-07-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: Custom
; SEQ ID NO 113
; LENGTH: 2731
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)...(2705)
US-10-188-186-113
Query Match 90.7%; Score 566.2; DB 17; Length 2731;
Best Local Similarity 97.3%; Pred. No. 8.7e-166;
Matches 582; Conservative 0; Mismatches 15; Indels 1; Gaps 1;
QY 27 CCTTTAGTNCAGATGACTTCCCATTCGCGAGGATCTTTGAATGTGTTCTTCTTGTTA 86
Db 1620 CCTTTAGTACAGATGACTTCCCATTCGCGAGGATCTTTGAATGTGTTCTTCTTGTTA 1561
QY 87 TGATCAACAAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGGCTTCTGGCAGG 146
Db 1560 TGATCAACAAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGGCTTCTGGCAGG 1501
QY 147 GGTTCCTTCAATCTGTAAAGCTTGAGTTCCTGGATCTTCCACCTGGGTTGTTTGTGTTA 206
Db 1500 GGTTCCTTCAATCTGTAAAGCTTGAGTTCCTGGATCTTCCACCTGGGTTGTTTGTGTTA 1441
QY 207 TGATTCACAAAGTAAACCTGCTGTGTGAATCCACTCTTTTCCAGGCTGTTGGCAAA 266
Db 1440 TGATTCACAAAGTAAACCTGCTGTGTGAATCCACTCTTTTCCAGGCTGTTGGCAAA 1381
QY 267 GGTTCATTAAGGTCATTTCTGAGCTAACATTTGAAGCCGAATAGAGGATCGTTGTTA 326
Db 1380 GGTTCATTAAGGTCATTTCTGAGCTAACATTTGAAGCCGAATAGAGGATCGTTGTTA 1321

```
QY 327 AACTGTGATAGTCTCCCTCAATTTGGTTCGGCTGAGATGGCCACTGNTCAAAATTNCGG 386
Db 1320 AACTGTGATAGTCTCCCTCAATTTGGTTCGGCTGAGATGGCCACTGTTTCAAAATTCGG 1261
QY 387 ACAGATTCCATGTAGGCGCTCCAGTGTGTTGTTCTGCTGTTATGATCCACATAATAA 446
Db 1260 ACAGATTCCATGTAGGCGCTCCAGTGTGTTGTTCTGCTGTTATGATCCACATAATAA 1201
QY 447 ACTCCTCTACGATCATCAACTCTTTTCCAACTGGAGGTAAGGTTGTTGCTCTCC 506
Db 1200 ACTCCTCTACGATCATCAACTCTTTTCCAACTGGAGGTAAGGTTGTTGCTCTCC 1141
QY 507 CATGTGTAGTTCGAGCATTTATGATCCACATAATAGCTCTACATGAGGATCTTTTCTT 566
Db 1140 CATGTGTAGTTCGAGCATTTATGATCCACATAATAGGTTTACATGAGGATCTTTTCTT 1081
QY 567 TGTTCCTCCACCTCGATGGCAAGNTTCTGTGTTGGCATTCCTCCAGACTGCTGCCGNACA 624
Db 1080 TGTTCCTCA-CCCTGATGGCAAGGTTTCTGTGTTGGCATTCCTCCAGACTGCTGCCGTACA 1024

RESULT 7
US-10-618-408-1/c
; Sequence 1, Application US/10618408
; Publication No. US20040191220A1
; GENERAL INFORMATION:
; APPLICANT: Paz Eilat, Louis Deiss, and Ruth Maya
; FILE REFERENCE: 67722-A; 074/US1
; CURRENT APPLICATION NUMBER: US/10/618,408
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: 60/395358
; PRIOR FILING DATE: 2002-07-11
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 2793
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2793)
US-10-618-408-1
```

```
Query Match 90.7%; Score 566.2; DB 19; Length 2793;
Best Local Similarity 97.3%; Pred. No. 8.8e-166;
Matches 582; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

QY 27 CCCTTTAGTNCAGATGACTCCCATTTGCGAGGATCTTTGAAATGTTGTTGTTGTTA 86
Db 1610 CCCTTTAGTTACAGATGACTTCCCATTTGCGAGGATCTTTGAAATGTTGTTGTTA 1551
QY 87 TGNATCAACAAGTACCTTACACCTTCCAGAGTATATCTAATTTCCAGCCCTCTGCGAGG 146
Db 1550 TGNATCAACAAGTACCTTACACCTTCCAGAGTATATCTAATTTCCAGCCCTCTGCGAGG 1491
QY 147 GGTTCCTTCATTCTGTAAAGCTTTCAGTTCCTGAGTCTTCCACCTGGGTTGTTTGTGTTA 206
Db 1490 GGTTCCTTCATTCTGTAAAGCTTTCAGTTCCTGAGTCTTCCACCTGGGTTGTTTGTGTTA 1431
QY 207 TGATTCACAAAGTAAACCTCTGTTGTAATCCACCTTTTTCAGCTGTTGTTGTTGTTGTTA 266
Db 1430 TGATTCACAAAGTAAACCTCTGTTGTAATCCACCTTTTTCAGCTGTTGTTGTTGTTGTTA 1371
QY 267 GGTTCATAGGGTCATTTTCTGAGCTTAACATTTGAGCCGAATGAGGATCTGTTGTTA 326
Db 1370 GGTTCATAGGGTCATTTTCTGAGCTTAACATTTGAGCCGAATGAGGATCTGTTGTTA 1311
QY 327 AACTGTGATAGTCTCCCTCAATTTGGTTCGGCTGAGATGGCCACTGNTCAAAATTNCGG 386
Db 1310 AACTGTGATAGTCTCCCTCAATTTGGTTCGGCTGAGATGGCCACTGNTCAAAATTNCGG 1251
QY 387 ACAGATTCCATGTAGGCGCTCCAGTGTGTTGTTCTGCTGTTATGATCCACATAATAA 446
```

```
Db 1250 ACAGATTCCATGTAGGCGCTCCAGTGTGTTGTTCTGCTGTTATGATCCACATAATAA 1191
QY 447 ACTCCTCTACGATCATCAACTCTTTTCCAACTGGAGGTAAGGTTGTTGCTCTCC 506
Db 1190 ACTCCTCTACGATCATCAACTCTTTTCCAACTGGAGGTAAGGTTGTTGCTCTCC 1131
QY 507 CATGTGTAGTTCGAGCATTTATGATCCACATAATAGCTCTACATGAGGATCTTTTCTT 566
Db 1130 CATGTGTAGTTCGAGCATTTATGATCCACATAATAGGTTTACATGAGGATCTTTTCTT 1071
QY 567 TGTTCCTCCACCTCGATGGCAAGNTTCTGTGTTGGCATTCCTCCAGACTGCTGCCGNACA 624
Db 1070 TGTTCCTCA-CCCTGATGGCAAGGTTTCTGTGTTGGCATTCCTCCAGACTGCTGCCGTACA 1014

RESULT 8
US-10-097-534-2/c
; Sequence 2, Application US/10097534
; Publication No. US20030049607A1
; GENERAL INFORMATION:
; APPLICANT: GREENER, TSVIKA
; APPLICANT: MOSKOWITZ, HAIM
; APPLICANT: REISS, YUVAL
; APPLICANT: ALROY, IRIS
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE MODULATION OF VIRAL
; FILE REFERENCE: PLV-001.01
; CURRENT APPLICATION NUMBER: US/10/097,534
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/275,224
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/308,958
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/340,170
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2377
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-097-534-2
```

```
Query Match 39.5%; Score 246.6; DB 14; Length 2377;
Best Local Similarity 64.3%; Pred. No. 6.6e-56;
Matches 388; Conservative 0; Mismatches 200; Indels 15; Gaps 1;

QY 12 TAAGCATTTGTGGACCCCTTTTAGTNCAGATGACTTCCCATTTGCGAGGATCTTTGAATGTT 71
Db 1079 TAGGCTATCTGAGGTCCATTGTCTAGGGCAGATTTTCTGTGCGGGGATCTATATAGGTG 1020
QY 72 GTTGTCTTGTGTTATGATCAACAAGTACCTTACACCTTCCAGAGTATATCTAATTTCC 131
Db 1019 GTAGTTCCTTCTATTTGTGGTCCAAATAATATGGAATTCCTGAAATCTCAATTTCC 960
QY 132 CAGCCTTCTGCGAGGGGTTCTTCATTCTGTAAAGCTTTCAGTTCCTGAGTCTTCCCATGG 191
Db 959 CAACCTTTCAGGTAAGGGCTTTTCATTAAATGACCTTTCAGTTCCTGAGGTTCTTCCCATGG 900
QY 192 GTTGTGTTTGTGTTATGATTCACAAAGTAAACCTGTCTGTTGAAATCCACTCTTTTTC 251
Db 899 GTAAATTCGTGTGTTGTTGTTGACGAAATATACCTGCTGCAATTCCTGTTCTCTTCTCC 840
QY 252 CAGCCTGTTGGCAAGGTCATTAAGGTCATTTTCTT-----GGAGCTAAC 296
Db 839 CATCCAGGTGGCAATGGACCAAGAGGATCAAAATCTTACTTTGTGATGAGCAATAATAA 780
QY 297 ATTGAAGCCGAATAGAGGATCGTTGGTTAAACTGTTGATAGTCTCCCTGCAATTTGTTTC 356
Db 779 TCTTGATTCCCATTAATGATCTCTGTTAAACTGTTGATGCTCTCTCTGAAAGCTGACTA 720
QY 357 CGCTGAGATTGCCACTGTGNTCAAAATTCGAGACAGATTCATGTTAGGCGCTGCCACGTT 416
```

Db 719 CCGTGTAGCTGCCATTGTTTCATAGTTCGGACGGATTCCAGTGTGGCTCTGCGCACGTT 660
QY 417 GTTGTCTGGGTATGATGCCATAAATAAACTCTCTACGATCAATCAACTCTTCTTTCC 476
Db 659 GTTGTCTTGTGAAATGGTCAACATAAATAAATACGTCCTCCATGTTGTCAACCGCGGTTCC 600
QY 477 CAACCTGGAGGTAAAGGTGTGGTCTCTCCCATGTGGTAGTTCGAGCATATGANGCCACA 536
Db 599 CAGCCAGGAGGTAGAGGTCTGGTCTATCCCAATGTTGTTCTTCTCAACATGATCTACA 540
QY 537 TAATACGCTCTACCATGAGGATCTTTTCTTTTGTCTCCACCCCTGATGGCAANGNTTCTGT 596
Db 539 TAGTAAACTCGCCGTCGTGGTCCACTCTCTGCTCCCAACAGGTGGCAAGGAGCTTGA 480
QY 597 GTT 599
Db 479 GTT 477

RESULT 9

US-10-374-979-65/c
; Sequence 65, Application US/10374979
; Publication No. US20030219793A1
; GENERAL INFORMATION:
; APPLICANT: John P. Carulli et al.
; TITLE OF INVENTION: THE HIGH BONE MASS GENE OF 11q13.3
; FILE REFERENCE: 032796-021
; CURRENT APPLICATION NUMBER: US/10/374,979
; CURRENT FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: US 09/544,398
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/543,771
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13
; PRIOR APPLICATION NUMBER: US 60/071,449
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/105,511
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 109
; SEQ ID NO 65
; LENGTH: 2377
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-374-979-65

Query Match 39.5%; Score 246.6; DB 17; Length 2377;
Best Local Similarity 64.3%; Pred. No. 6.6e-66;
Matches 388; Conservative 0; Mismatches 200; Indels 15; Gaps 1;
QY 12 TAAGCATTTGTGACCCCTTTAGTNCAGATGACTCCCATTCGGAGGATCTTTGAATGTT 71
Db 1079 TAGGCTATCTGAGGTCCATTGTTCTAGGCGAGATTTTCCTGTGGGGGATCTATATAGGTG 1020
QY 72 GTTGTCTTCTGTGTATGATCAACAAAGTACTTTACACCTTCACGAGTATATCTAATTTCC 131
Db 1019 GTAGTCTTCTTATTTGGTCCACAAATATGGAATTCATCCACTGTGAATCTCAATTC 960
QY 132 CAGCCTTCGGAGGGGTCTTCATTTCTGTAAAGCTTTGAGTCTTTCGATCTTCCACTGG 191
Db 959 CAACCTTCAGTAAAGGCTTTTCAATTAATGACCTTGATCTTGGGGTCTTCCCAATGT 900
QY 192 GTTGTCTTCTGTGTATGATTCACAAAGTAAACCTGTGTGTAATCCACTCTTTTTC 251
Db 899 GTAAATCGTGTGTGTGATGATTCACAAAGTAAACCTGTGTGTAATCCACTCTTTTTC 840
QY 252 CAGCCTTCGGAGGGTCCATTAAGGCTCAATTTCT-----GCAGTAAAC 296
Db 839 CATCCAGGTGGCAATGACCAAGAGGATCAAAATCTTTACTTGTGTGATGTAGCAATAA 780
QY 297 ATTGAAGCCGAATAGAGGTATCGTTGGTTAAACTGTGTGATAGTCCCTGCAATTTGTT 356

Db 779 TCTTGATTTCCCATAAATGAATCTCTGGTTAAACTGCTGCATTTGCTCTTGAAGTGAATA 720
QY 357 CGCTGAGATTGCCACTGNTCAAAATTTNCGGACAGATTCATGTTAGGCGGCTGCGCACGTT 416
Db 719 CCGTGTAGCTGCCATTGTTTCATAGTTCGGACGGATTCAGTGTGGCTCTGCGCACGTT 660
QY 417 GTTGTCTTGGTATGATCCACATAAATAAACTCTCTACGATCAATCAACTCTTCTTTCC 476
Db 659 GTTGTCTTGTGAAATGGTCAACATAAATAAATACGTCCTCCATGTTGTCAACCGCGGTTCC 600
QY 477 CAACCTGGAGGTAAAGGTGTGGTCTCTCCCATGTGGTAGTTCGAGCATATGANGCCACA 536
Db 599 CAGCCAGGAGGTAGAGGTCTGGTCTATCCCAATGTTGTTCTTCTCAACATGATCTACA 540
QY 537 TAATACGCTCTACCATGAGGATCTTTTCTTTTGTCTCCACCCCTGATGGCAANGNTTCTGT 596
Db 539 TAGTAAACTCGCCGTCGTGGTCCACTCTCTGCTCCCAACAGGTGGCAAGGAGCTTGA 480
QY 597 GTT 599
Db 479 GTT 477

RESULT 10

US-10-240-425-383/c
; Sequence 383, Application US/10240425
; Publication No. US20040033502A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Amanda
; APPLICANT: Boland, Joseph F.
; APPLICANT: Lord, Reginald V.
; APPLICANT: Alvarez, Chris
; APPLICANT: Wetzel, Jon C.
; APPLICANT: Scherf, Uwe
; APPLICANT: Vockley, Joseph G.
; TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue
; FILE REFERENCE: 44921-5026
; CURRENT APPLICATION NUMBER: US/10/240,425
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: PCT/US01/09847
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 60/193,446
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 1588
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 383
; LENGTH: 2377
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20040033502A1 AF038564
US-10-240-425-383

Query Match 39.5%; Score 246.6; DB 18; Length 2377;
Best Local Similarity 64.3%; Pred. No. 6.6e-66;
Matches 388; Conservative 0; Mismatches 200; Indels 15; Gaps 1;
QY 12 TAAGCATTTGTGAGCCCTTTAGTNCAGATGACTTTCCCATTCGGAGGATCTTTGAATGTT 71
Db 1079 TAGGCTATCTGAGGTCCATTGTTCTAGGCGAGATTTTCCTGTGGGGATCTATATAGGTG 1020
QY 72 GTTGTCTTGTGTATGATCAACAAAGTACTTTACACCTTCACGAGTATATCTAATTTCC 131
Db 1019 GTAGTCTTCTTATTTGGTCCACAAATATGGAATTCATCCACTGTGAATCTCAATTTCC 960
QY 132 CAGCCTTCGGAGGGGTCTTTCATTTCTGTAAAGCTTTGAGTCTTTCGATCTTCCCACTGG 191
Db 959 CAACCTTCAGTAAAGGCTTTTCAATTAATGACCTTGATCTTGGGGTCTTCCCAATGT 900
QY 192 GTTGTCTTGTGTATGATTCACAAAGTAAACCTGTCTGTGTAATCCACTCTTTTTC 251
Db 899 GTAAATCGTGTGTGTGATGATTCACAAAGTAAATACTCTGCCATTTGCTGTCTTCTTCC 840

Qy	252	CAGCCTGTGTGGCAAGGTCCATAAGGGTCATTTTCT-----GCAGCTAAC	296
Db	839	CATCAGGTGGCAATGGACCAAGAGGATCAAAATCTTTACTTTGTGATGATCAAAATAA	780
Qy	297	ATTGAAGCCGAATAGAGGTCGTGTGGTTTAAACTGTTCATAGCTCCCTGCAATTTGGTTC	356
Db	779	TCTTGATTTCCCATAAATGAATCTGTGGTTTAAACTGCTGCAATTTGCTCTTTGAAGCTGACTA	720
Qy	357	CGCTGAGATTCGCCACTGNTCAAATTTNCGGACAGAGATTCCATGTTAGGGCCGCTGCCACGTT	416
Db	719	CGCTGTAGCTGCCATTGTTTCATAGTTCCGGACGGAATCCAGTGTGGCCCTCTGCCACGTT	660
Qy	417	GTTGTTCTGTGGTTTATGATCCACATAATAAATCTCTTACGATCATCAACTCTTCTTTCC	476
Db	659	GTTGTTCTTGTGAAATGGTCAACATAATAAATACGTCCCATGTTGTTCACACCGCCGCTTCC	600
Qy	477	CAACCTGGAGGTAAAGGTTGTGGTCTCTCCCATGTGGTAGTTTCGAGCATATGAMCCACA	536
Db	599	CAGCCAGAGGTAGAGGTTCTGGTCTATCCCATGTGTCTTTCTTCAACATGATCTACA	540
Qy	537	TAATACGCTTACCATGAGGATCTTTTCTTTTGTCCCACTCCCTGATGGCAANGNTTCTGT	596
Db	539	TAGTAAACTCCCGCTGTGGTCCACTCTCTGCTCCCAACCAAGGTGGCAAGGAGCTTGA	480
Qy	597	GTT	599
Db	479	GTT	477

RESULT 11

```

US-10-182-936A-65/c
; Sequence 65, Application US/10182936A
; Publication No. US20040038860A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Kristina M.
; APPLICANT: Anisowicz, Anthony
; APPLICANT: Bhat, Bheem
; APPLICANT: Damagnez, Veronique
; APPLICANT: Robinson, John
; APPLICANT: Yaworsky, Paul
; TITLE OF INVENTION: Reagents and Method for Modulating DKK-Mediated Interactions
; FILE REFERENCE: 032796-143
; CURRENT APPLICATION NUMBER: US/10/182,936A
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: PCT/US02/15982
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 65
; LENGTH: 2377
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-182-936A-65

```

[illegible]

RESULT 12

```

US-10-283-975A-128/c
; Sequence 128, Application US/10283975A
; Publication No. US20040110792A1
; GENERAL INFORMATION:
; APPLICANT: Ortho-Clinical Diagnostics, Inc.
; TITLE OF INVENTION: Methods For Assessing and Treating Leukemia
; FILE REFERENCE: CDS 293 PCT
; CURRENT APPLICATION NUMBER: US/10/283,975A
; CURRENT FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/340,938
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/338,997
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/340,081
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/341,012
; PRIOR FILING DATE: 2001-10-30
; NUMBER OF SEQ ID NOS: 900
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 128
; LENGTH: 2377
; TYPE: DNA
; ORGANISM: HUMAN
US-10-283-975A-128

```

	Query Match	39.5%	Score 246.6	DB 18	Length 2377	
	Best Local Similarity	64.3%	Pred. No. 6.6e-66			
	Matches 388	Conservative 0	Mismatches 200	Indels 15	Gaps 1	
Qy	12	TAAGCATTTGTGGACCCCTTTAGTNCAGATGACTTCCCATTTGGAGAGATCTTTGAATGTT	71			
Db	1079	TAGGCTATCTGAGGTCCATTTGTCTAGGCGCAGATTTTCTCTGTGGGGGATCTATATAGTG	1020			
Qy	72	GTTCGTTCTTGTTTATGATCAACAAAGTCACTTACACTTCACAGATATATCTAAATTTCC	131			
Db	1019	GTATGTTCTTCTATTGTGGTCCAAAATATGGAAATTCATCCACTGTGAAATCTCATTTCC	960			
Qy	132	CAGCCTTCTCGCAGGGGTCTTTCATCTGTAAGCCTTTGAGTTCCTTGATCTTCCCACTGG	191			

```
QY 132 CAGCCTTTCGGCAGGGGTTCTTCAATCTGTGAAGCCTTGAGTTCTTGGATCTTCCCACTGG 191
|||
Db 959 CAACCTTCAGGTAAAGGCTTTTCAATTAATTAATGACCTTGACTCTGGGGTCTTCCCAATGT 900
|||
QY 192 GTTGTGTTTGTGTATGATTCACAAAGTAAACCCCTGTCTGTGTAATCCCACTCTTTTTC 251
|||
Db 899 GTAAATTCGTGTGTGTGTGAGGAAATATACTCTGCCATTTGCTGTCTTCTTCTCTCC 840
|||
QY 252 CAGCCTGGTGGCAAGGTCATTAAGGTCATTTTCT - - - - -GCAGCTAAC 296
|||
Db 839 CATCCAGGTGGCAATGGACCAAGAGGATCAAAATCTTTTGTGATGAGCAAAATAA 780
|||
QY 297 ATTGAAGCCGAATAGAGGTATCGTTGTTAAATCTGTTCATAGCTCCCTGCAATTCGGTTC 356
|||
Db 779 TCTTGATTCCTCAATAATGAATCTCTGTTTAAACTGCTGCAATGCTCTTGAAGCTGACTA 720
|||
QY 357 CGCTGAGATTGCCACTGTNTCAAAATTCGACAGATTCATGATGAGGCGCTGCCACGTT 416
|||
Db 719 CGCTGTAGCTGCCATGTTTCAATAGTTCGAGCAGGATTCAGTGTGGCCTCTGCCACGTT 660
|||
QY 417 GTTGTCTCGTGTATGATCCACATAATAAACTCCTCTACGATCATCAACTCTTCTTCC 476
|||
Db 659 GTTGTCTTGTGAATGGTCAACATAATAAATACGTCATGTTGTCAACCCGCGTTC 600
|||
QY 537 TAATACGCTTACCATGAGATCTTTTGTGTTCCACCCCTGTAGTGGCAANGTTCTGT 596
|||
Db 539 TAGTAACTCGCCGCTGTGTCCTCTCTGCTCCCAACCAAGGTGGCAAGGAGCTTGA 480
|||
QY 597 GTT 599
|||
Db 479 GTT 477

RESULT 13
US-10-477-238A-644/c
; Sequence 644, Application US/10477238A
; Publication No. US20040221326A1
; GENERAL INFORMATION:
; APPLICANT: Babij, Philip
; APPLICANT: Yaworsky, Paul
; APPLICANT: Bex, Frederick J. III
; APPLICANT: Bodine, Peter Van Nest
; TITLE OF INVENTION: Transgenic Animal Model of Bone Mass Modulation
; FILE REFERENCE: 032796-212
; CURRENT APPLICATION NUMBER: US/10/477,238A
; PRIOR FILING DATE: 2003-11-10
; PRIOR APPLICATION NUMBER: US 60/290,071
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 812
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 644
; LENGTH: 2377
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-477-238A-644

Query Match 39.5%; Score 246.6; DB 20; Length 2377;
Best Local Similarity 64.3%; Pred. No. 6.6e-56;
Matches 388; Conservative 0; Mismatches 200; Indels 15; Gaps 1;

QY 12 TAAGCATTTGTGAGCCCTTTAGTCAGATGCTTCCCATTTGGAGATCTTTGAATGTT 71
|||
```

```
Query Match 39.5%; Score 246.6; DB 20; Length 2377;
Best Local Similarity 64.3%; Pred. No. 6.6e-66;
Matches 388; Conservative 0; Mismatches 200; Indels 15; Gaps 1;

QY 12 TAAGCATTTGGACCCCTTTAGTNCAGATGACTTCCCATTTGGAGGATCTTTGAATGTT 71
DB 1079 TAGGCTATCTGAGGTCCATTTGCTTAGGGCAGATTTTCTGTGGGGGATCTATATAGTG 1020

QY 72 GTTGTCTTCTGTGTATGATCAACAAAGTACCTTACACCTTCCAGAGTATATCTAAATTC 131
DB 1019 GTAGTCTTCTATTTGGTCCACAAATATGAATTCATCCACCTGTGAATCTCAATTC 960

QY 132 CAGCTTCTGGCAGGGGTTCTTCAATCTGTAAGCCCTTGAAGTCTTGTGATCTTCCACTGG 191
DB 959 CAACCTTCAGTAAGGGCTTTTCAATTAATGACCTTGACTTCTGGGGTCTTCCATTTG 900

QY 192 GTTGTCTTCTGTGTATGATCAACAAAGTACCTTACACCTTCCAGAGTATATCTAAATTC 251
DB 899 GTAAATTCGTTGTGTGTGACGAAATATGTAATCTGCAATGCTGCTGTCTCTCTCC 840

QY 252 CAGCTTCTGGCAGGGGTTCTTCAATCTGTAAGCCCTTGAAGTCTTGTGATCTTCCACTGG 296
DB 839 CATCCAGGTGCAATGACCAAGAGGATCAAAATCTTTACTTTGTGATGTAGCAATATA 780

QY 297 ATTGAAGCCGAATAGAGGTATGCTGTGTTAACTGTTGCAATGCTCCCTGCAATTTGTT 356
DB 779 TCTGTATTCCTCAATTAATGAATCTCTGTTAACTGTTGCAATGCTCTTGAAGTACTA 720

QY 357 CGCTGAGATTCGCACTGNTCAAAATTCGAGAGGATTCATGAGTCTTCCACTCTTCTTCC 416
DB 719 CGCTGAGTCTGCACTGNTCAAAATTCGAGAGGATTCATGAGTCTTCCACTCTTCTTCC 660

QY 417 GTTGTCTTCTGTGTATGATCAACAAAGTACCTTACACCTTCCAGAGTATATCTAAATTC 476
DB 539 TAGTAAACTCGCCGCTGCTGCTCCACTCTCTGTCTCCCAACCAAGTGGCAAGGAGCTTGA 480

QY 597 GTT 599
DB 479 GTT 477
```

RESULT 15
US-10-477-173-644/c
; Sequence 644, Application US/10477173
; Publication No. US20050070699A1
; GENERAL INFORMATION:
; APPLICANT: Genome Therapeutics Corporation and
; APPLICANT: Allen, Kristina M.
; APPLICANT: Yaworsky, Paul
; APPLICANT: Morales, Arturo J.
; APPLICANT: Graham, James R.
; APPLICANT: Anisowicz, Anthony
; APPLICANT: Liu, Wei
; TITLE OF INVENTION: HBM Variants that Modulate Bone Mass and Lipid Levels
; FILE REFERENCE: 032796-135
; CURRENT APPLICATION NUMBER: US/10/477,173
; CURRENT FILING DATE: 2003-11-10
; PRIOR APPLICATION NUMBER: US 60/290,071
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293

```
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 1086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 644
; LENGTH: 2377
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-477-173-644

Query Match 39.5%; Score 246.6; DB 21; Length 2377;
Best Local Similarity 64.3%; Pred. No. 6.6e-66;
Matches 388; Conservative 0; Mismatches 200; Indels 15; Gaps 1;

QY 12 TAAGCATTTGTGACCCCTTTAGTNCAGATGACTTCCCATTTGGAGGATCTTTGAATGTT 71
DB 1079 TAGGCTATCTGAGGTCCATTTGCTTAGGGCAGATTTTCTGTGGGGGATCTATATAGTG 1020

QY 72 GTTGTCTTCTGTGTATGATCAACAAAGTACCTTACACCTTCCAGAGTATATCTAAATTC 131
DB 1019 GTAGTCTTCTATTTGGTCCACAAATATGAATTCATCCACCTGTGAATCTCAATTC 960

QY 132 CAGCTTCTGGCAGGGGTTCTTCAATCTGTAAGCCCTTGAAGTCTTGTGATCTTCCACTGG 191
DB 959 CAACCTTCAGTAAGGGCTTTTCAATTAATGACCTTGACTTCTGGGGTCTTCCATTTG 900

QY 192 GTTGTCTTCTGTGTATGATCAACAAAGTACCTTACACCTTCCAGAGTATATCTAAATTC 251
DB 899 GTAAATTCGTTGTGTGTGACGAAATATGTAATCTGCAATGCTGCTGTCTCTCTCC 840

QY 252 CAGCTTCTGGCAGGGGTTCTTCAATCTGTAAGCCCTTGAAGTCTTGTGATCTTCCACTGG 296
DB 839 CATCCAGGTGCAATGACCAAGAGGATCAAAATCTTTACTTTGTGATGTAGCAATATA 780

QY 297 ATTGAAGCCGAATAGAGGTATGCTGTGTTAACTGTTGCAATGCTCCCTGCAATTTGTT 356
DB 779 TCTGTATTCCTCAATTAATGAATCTCTGTTAACTGTTGCAATGCTCTTGAAGTACTA 720

QY 357 CGCTGAGATTCGCACTGNTCAAAATTCGAGAGGATTCATGAGTCTTCCACTCTTCTTCC 416
DB 719 CGCTGAGTCTGCACTGNTCAAAATTCGAGAGGATTCATGAGTCTTCCACTCTTCTTCC 660

QY 417 GTTGTCTTCTGTGTATGATCAACAAAGTACCTTACACCTTCCAGAGTATATCTAAATTC 476
DB 659 GTTGTCTTCTGTGAAATGTTGTCACATATAAATACGTCCTCCATGTTGTTCACCCGCTTCC 600

QY 477 CAACCTGGAGGTAAAGTGTGTTGCTCTCCCATGTGTTGAGTCTGAGCATTTATGANCACA 536
DB 599 CAGCCAGGAGGTAGAGTCTGCTCTATCCCATGTTGTTCTTTCTCAACATGATCTACA 540

QY 537 TAATACGCTCTACCATGAGGATCTTTTCTTTTGTTCACCCCTGATGGCAANGTTCGT 596
DB 539 TAGTAAACTCGCCGCTGCTGCTCCACTCTCTGTCTCCCAACCAAGTGGCAAGGAGCTTGA 480

QY 597 GTT 599
DB 479 GTT 477
```

Search completed: June 30, 2005, 06:39:42
Job time : 634 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2005, 04:55:15 ; Search time 2866 Seconds
(without alignments)
8287.545 Million cell updates/sec

Title: US-10-618-408-3
Perfect score: 624
Sequence: 1 ccgcgcgtnttaagcattt.....tccagactgctgcgcgnaca 624

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 69479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_ges1.*
9: gb_ges2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	525.2	84.2	899	7	CF551931
C 2	508.6	81.5	3107	3	BC026829
C 3	508.6	81.5	4628	3	AK033138
C 4	495.6	79.4	742	7	CN533573
C 5	487.6	78.1	3823	3	AK032232
C 6	481.8	77.2	1113	7	CK233395
C 7	473.4	75.9	654	7	CK981350
C 8	453.4	72.7	655	2	BB624213
C 9	434.8	69.7	1101	2	BE869796
C 10	425.8	68.2	620	4	BI598894
C 11	416	66.7	891	5	BU050208
C 12	391.6	62.8	630	1	AL049041
C 13	380.6	61.0	655	6	CB297052
C 14	376	60.3	474	1	AI750898
C 15	373.6	59.9	557	2	BF369480
C 16	367.4	58.9	527	2	BF333137
C 17	341.2	54.7	448	5	BQ304314
C 18	339.2	54.4	469	2	BF333147
C 19	333	53.4	519	2	BF333178
C 20	329.2	52.8	731	7	CK470232
C 21	322.8	51.7	394	5	BQ304311
C 22	317.2	50.8	810	7	CK474855
C 23	309	49.5	762	6	CD802832
C 24	307.8	49.3	401	2	BF333140

25	307.6	49.3	470	2	BF333183
26	300.2	48.1	445	5	BQ365503
C 27	292.4	46.9	899	7	CF583889
C 28	287.6	46.1	569	4	BG969441
C 29	286.8	46.0	687	4	BI459570
C 30	283.4	45.4	814	2	BF119329
C 31	276.6	44.3	512	4	BI713321
C 32	271.6	43.5	563	4	BI884127
C 33	260.4	41.7	712	6	CB457326
C 34	260	41.7	854	7	CV117296
C 35	251	40.2	692	7	CF747411
C 36	249.2	39.9	716	7	CO432363
C 37	249.2	39.9	1774	3	AK037940
C 38	249.2	39.9	2481	3	AK048303
C 39	246.6	39.5	1050	5	BM917937
C 40	244.2	39.1	704	7	CN397454
C 41	240.8	38.6	528	1	AJ727321
C 42	239	38.3	582	5	BP221259
C 43	238.2	38.2	674	1	AJ727320
C 44	232.2	37.2	700	7	CK635728
C 45	231.6	37.1	581	5	BQ379860

ALIGNMENTS

RESULT 1
CF551931/c
LOCUS
DEFINITION AGENCOURT 15595822 NIH_MGC_183 Homo sapiens cDNA clone
IMAGE:30530049 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CF551931 899 bp mRNA linear EST 22-SEP-2003
AGENCOURT 15595822 NIH_MGC_183 Homo sapiens cDNA clone
IMAGE:30530049 5', mRNA sequence.
CF551931
CF551931.1 GI:34888765
EST.
Homo sapiens (human)
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: csapbe-x@mail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Invitrogen Corp
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: Agencourt Bioscience Corporation
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM618 row: j column: 10
High quality sequence start: 90
High quality sequence stop: 700.
Location/Qualifiers
1..899
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30530049"
/lab_host="DH10B-Tona (T1 and T5 phage resistant)"
/clone_lib="NIH_MGC_183"
/note="Organ: Pooled muscle (cardiac and skeletal);
Vector: pCMV-SPORT6.1; Site 1: EcoRV (destroyed); Site 2:
NotI; Library is oligo-dr primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 1.7. Library was constructed by Invitrogen."

ORIGIN

Query Match 84.2%; Score 525.2; DB 7; Length 899;

Best Local Similarity 96.8%; Pred. No. 5e-144; Matches 544; Conservative 0; Mismatches 17; Indels 1; Gaps 1;	
QY 27	CCCTTTAGTNCAGATGACTTCCCATTCGAGGATCTTTGAATGTTGTTGTTCTTGTGTTA 86
Db	577 CCTTTAGTTCAGATGACTTCCCATTCGAGGATCTTTGAATGTTGTTGTTCTTGTGTTA 518
QY 87	TGATCAACAAGTACTTTACACCTTCACGAGTATATCTAAATTCACGACCTTCGCGAGG 146
Db	517 TGATCAACAAGTACTTTACACCTTCACGAGTATATCTAAATTCACGACCTTCGCGAGG 458
QY 147	GTTCTTCATCTGTAAGCCTTGAGTCTTCCCATTCGAGTATATCTAAATTCACGACCTTCGCGAGG 206
Db	457 GTTCTTCATCTGTAAGCCTTGAGTCTTCCCATTCGAGTATATCTAAATTCACGACCTTCGCGAGG 398
QY 207	TGATTCACAAAGTAAACCTTCGTTGTAATCCACTCTTTTCCAGCCTTCGCGAGG 266
Db	397 TGATTCACAAAGTAAACCTTCGTTGTAATCCACTCTTTTCCAGCCTTCGCGAGG 338
QY 267	GTTCCATAAGGGTCAATTTCTGCACTTAAACATTCGAGTATATCTAAATTCACGACCTTCGCGAGG 326
Db	337 GTTCCATAAGGGTCAATTTCTGCACTTAAACATTCGAGTATATCTAAATTCACGACCTTCGCGAGG 278
QY 327	AACGTGTCATAGCTCCCTGCAATTTGTTGCGTTCGAGTATATCTAAATTCACGACCTTCGCGAGG 386
Db	277 AACGTGTCATAGCTCCCTGCAATTTGTTGCGTTCGAGTATATCTAAATTCACGACCTTCGCGAGG 218
QY 387	ACAGATTCATAGTTCGCGTTCGCAATTTGTTGCGTTCGAGTATATCTAAATTCACGACCTTCGCGAGG 446
Db	217 ACAGATTCATAGTTCGCGTTCGCAATTTGTTGCGTTCGAGTATATCTAAATTCACGACCTTCGCGAGG 158
QY 447	ACTCCTTCATAGTTCGCGTTCGCAATTTGTTGCGTTCGAGTATATCTAAATTCACGACCTTCGCGAGG 506
Db	157 ACTCCTTCATAGTTCGCGTTCGCAATTTGTTGCGTTCGAGTATATCTAAATTCACGACCTTCGCGAGG 98
QY 507	CATGTG-GTAGTTCGAGTATATGACATTCATATACGCTCTTACCATGAGGATCTTTTCT 565
Db	97 CATGTGNTAGTTCGAGTATATGACATTCATATACGCTCTTACCATGAGGATCTTTTCT 38
QY 566	TTGTTTCCACCCCTTGATGGCAA 587
Db	37 TTGTTCCACCCCTTGATGGCAA 16
RESULT 2	
BC026829/c	
LOCUS	
DEFINITION Mus musculus WW domain containing E3 ubiquitin protein ligase 1, mRNA (CDNA clone IMAGE:4209397), containing frame-shift errors.	
ACCESSION BC026829.1	
VERSION GI:20071294	
KEYWORDS	
SOURCE Mus musculus (house mouse)	
ORGANISM	
REFERENCE	
AUTHORS	
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalilus, D.E., Schnerch, A., Schein, J.E., Jones, S.O. and Marra, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)	
2 (bases 1 to 3107)	
Strausberg, R. Direct Submission Submitted (04-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
NIH-MGC Project URL: http://mgc.nci.nih.gov	
Contact: MGC help desk Email: cgaps-i@mail.nih.gov Tissue Procurement: Jeffrey E. Green, M.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: http://www.hgsc.bcm.tmc.edu/cdna/ Contact: amg@bcm.tmc.edu Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsegad, H., Kowis, C.R., Speed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.	
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAK Plate: 37 Row: h Column: 15 This clone has the following problem: frame shifted.	
FEATURES	
source	
1..3107	
/organism="Mus musculus"	
/mol_type="mRNA"	
/strain="FVB/N"	
/db_xref="taxon:10090"	
/clone="IMAGE:4209397"	
/tissue_type="Colon, normal. 5 month old male mouse."	
/clone_lib="NCI CGAP Co24"	
/lab_host="DH10B"	
/note="Vector: pCMV-SPORT6"	
ORIGIN	
Query Match 81.5%; Score 508.6; DB 3; Length 3107;	
Best Local Similarity 91.3%; Pred. No. 5.6e-139; Matches 546; Conservative 0; Mismatches 51; Indels 1; Gaps 1;	
QY 27	CCCTTTAGTNCAGATGACTTCCCATTCGAGGATCTTTGAATGTTGTTGTTCTTGTGTTA 86
Db	1657 CCTTTAGTTCAGATGACTTCCCATTCGAGGATCTTTGAATGTTGTTGTTCTTGTGTTA 1598
QY 87	TGATCAACAAGTACTTTACACCTTCACGAGTATATCTAAATTCACGACCTTCGCGAGG 146
Db	1597 TGATCAACAAGTACTTTACACCTTCACGAGTATATCTAAATTCACGACCTTCGCGAGG 1538
QY 147	GTTCTTCATCTGTAAGCCTTGAGTCTTCCCATTCGAGTATATCTAAATTCACGACCTTCGCGAGG 206
Db	1537 GTTCTTCATCTGTAAGCCTTGAGTCTTCCCATTCGAGTATATCTAAATTCACGACCTTCGCGAGG 1478
QY 207	TGATTCACAAAGTAAACCTTCGTTGTAATCCACTCTTTTCCAGCCTTCGCGAGG 266
Db	1477 TGATTCACAAAGTAAACCTTCGTTGTAATCCACTCTTTTCCAGCCTTCGCGAGG 1418
QY 267	GTTCCATAAGGGTCAATTTCTGCACTTAAACATTCGAGTATATCTAAATTCACGACCTTCGCGAGG 326
Db	1417 GTTCCATAAGGGTCAATTTCTGCACTTAAACATTCGAGTATATCTAAATTCACGACCTTCGCGAGG 1358
QY 327	AACTGTCATAGCTCCCTGCAATTTGTTGCGTTCGAGTATATCTAAATTCACGACCTTCGCGAGG 386
Db	1357 AACTGTCATAGCTCCCTGCAATTTGTTGCGTTCGAGTATATCTAAATTCACGACCTTCGCGAGG 1298

207 TGATTCACAAAGTAAACCCCTGCTGTTGAATCCACTCTTTTTCCAGCCTGGTGGCAA 266
|||||
Db TGATTCACAAAGTAAACCTGCTGTTGAATCCACTCTTTTTCCAGCCTGGTGGTAAG 456
|||||
Qy GGTCCATTAAGGGTCATTTTCTGAGCTAAATGTAAGCCCAATAGAGGTATCGTTGGTTA 326
|||||
Db GGTCCATTAAGGGTCATTTTCTGAGCTAAATGTAAGCCCAATAGAGGTATCGTTGGTTG 396
|||||
Qy AACTGTTGTCATAGCTCCCTGCAATTTGGTTCGGCTGAGATTGCCACTGNTCAAAATTCGG 386
|||||
Db AATTGTTGTCATAGCTCCCTGCAATTTGGTTCGGCTGAGATTGCCACTGNTCAAAATTCGA 336
|||||
Qy ACAGATTCCATGGTAGGCCCTGCCAGTGTGTTGTTCTGTTGTTATGATCAATATAA 446
|||||
Db ACAGATTCCATGGTAGGCCCTGCCAGTGTGTTGTTCTGTTGTTATGATCAATATAA 276
|||||
Qy ACTCCTCTAGCATCATCACTCTTTTCCCACTGGAGGTAAAGTTGTGTCTCTCC 506
|||||
Db ACTCCTCTAGCATCATCACTCTTTTCCCACTGGAGGTAAAGTTGTGTCTCTCC 216
|||||
Qy CATGTGTGTAGTTGAGCATTTATGANCACATAATAGCTCTACCATGAGGATCTTTTCTT 566
|||||
Db CATGTGTGTAGTTGAGCATTTATGATCCACATAATAGGTTCTACATGAGGATCTTTTCTC 156
|||||
Qy TGTTCACACCCCTGATGGCAANGTTCTGTGTTGGCATTTCCAGACTGCTGCCGNACA 624
|||||
Db TGTTCACCA-CCCTGATGGCAAGCTTCTGTGTTGGTGTTCAGACTGTGGCCGCAGA 99
|||||

RESULT 4

CN533573/c
LOCUS
DEFINITION
UT-M-HO0-CPY-P-14-0-UI.r1 NIH BMAP_H00 Mus musculus cDNA clone
IMAGE:30658453 5', mRNA sequence.
CN533573
EST.
CN533573.1 GI:46861729
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 742)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@email.nih.gov
Tissue Procurement: Dr. James Lin University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pYX-5.

FEATURES

source	Location/Qualifiers
	1..742
	/organism="Mus musculus"
	/mol_type="mRNA"
	/strain="C57BL/6"
	/db_xref="taxon:10090"
	/clone="IMAG:30658453"
	/tissue_type="Upper Head"
	/dev_stage="9.5-10.5 dpc"
	/lab_host="DH10B (T1 phage"
	/notes="NIH BMAP H00"
	/notes="Organ: Head; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was

AK032232 3823 bp mRNA linear HTC 03-APR-2004
Mus musculus adult male olfactory brain cDNA, RIKEN full-length
enriched library, clone:6430508G12 product:WMP1 (FRAGMENT) homolog
[Homo sapiens], full insert sequence.
AK032232
HTC; CAP trapper.
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci P. and Havaashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253

size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is CGAAGTGAAT. This library was created for the University
of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH)."

ORIGIN

Query Match	79.4%; Score 495.6; DB 7; Length 742;
Best Local Similarity	90.8%; Pred. No. 2.7e-135;
Matches 544; Conservative 0; Mismatches 53; Indels 2; Gaps 2;	
Qy 27 CCCTTTAGTNCAGATGACATTCCTCCATTCGCGAGATCTTTGAATGTTCTGTTCTGTGTTA 86 	
Db 639 CCTTTAGTTACTGATGACATTCCTCCATTCGCGAGGTCTNTGAAATGTTGTTCTCGGTAN 580 	
Qy 87 TGATCAACAAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGCCTTCTCGCAGG 146 	
Db 579 TGATCAACAAAGTACCTTACACCTTCCTCCGAGTATACCTGATTTCCAGCCTTCGGCAGG 520 	
Qy 147 GGTTCCTCATCTGTAAGCTTGAGTTCCTTGAGATCTTCCACTGGGTTGTTTGTGTTA 206 	
Db 519 GGTTCCTCATCTGTAAGCTTCGAGTTCCTTGAGATCTTCCACTGGGTTGTTTGTATTG 460 	
Qy 207 TGATTCACAAAGTAAACCCCTGCTGTTGAATCCACTCTTTTTTCCAGCCTGGTGGCAA 266 	
Db 459 TGATTCACAAAGTAAACCTGCTGTTGAATCCACTCTTTTTTCCAGCCTGGTGGTAAG 400 	
Qy 267 GGTTCATAAGGGTCAATTTTCTGCAGCTAACATTTGAAGCCGAATAGAGGTATCGTTGTTA 326 	
Db 399 GGTTCATAAGGGTCAATTTTCTGCAGCTAACATTTGAAGCCGAATAGAGGTATCGTTGTTG 340 	
Qy 327 AACTGTTGATAGCTCCCTGCAATTTGGTTCGGTGAATGCGCATGTCATGTCATCAAAATTCGG 386 	
Db 339 AATTGTTGATAGCTCCCTGCAATTTGGTTCGGTGAATGCGCATGTCATGTCATCAAAATTCGA 280 	
Qy 387 ACAGATTCCATGTTAGCGCTGCCAGCTGTTGTTCTGTTGTTATGATCCACATAATAA 446 	
Db 279 ACAGATTCCATGTTAGCGCTGCCAGCTGTTGTTCTGTTGTTATGATCCACATAATAA 220 	
Qy 447 ACTCCTCTAGCATCATCACTCTTTTCCCACTGGAGGTAAAGTTGTGTCTCTCC 506 	
Db 219 ACTCCTCTAGCATCATCACTCTTTTCCCACTGGAGGTAAAGTTGTGTCTCTCC 160 	
Qy 507 CATGTGTGTAGTTGAGCATTTATGANCACATAATAGCTCTACCATGAGGATCTTTTCTT 566 	
Db 159 CATGTGTGTAGTTGAGCATTTATGATCCACATAATAGGTTCTACCATGAGGATCTTTTCTC 100 	
Qy 567 TGTTCACACCCCTGATGGCAANGTTCTGTGTTGG-CAITTCACAGACTGCTGCCGNACA 624 	
Db 99 TGTTCACCA-CCCTGATGGCAAGCTTCTGTGTTGGTGTTCACAGACTGTGGCCGCAGA 42 	

RESULT 5

AK032232/c
LOCUS
DEFINITION
Mus musculus adult male olfactory brain cDNA, RIKEN full-length
enriched library, clone:6430508G12 product:WMP1 (FRAGMENT) homolog
[Homo sapiens], full insert sequence.
AK032232
HTC; CAP trapper.
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci P. and Havaashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253

10349636
2
REFERENCE
AUTHORS
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL
MEDLINE
20499374
PUBMED
11042159
REFERENCE
AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Taehiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL
MEDLINE
20530913
PUBMED
11076861
REFERENCE
AUTHORS
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE
Functional annotation of a full-length mouse cDNA collection
JOURNAL
MEDLINE
11076861
PUBMED
11076861
REFERENCE
AUTHORS
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL
MEDLINE
20530913
PUBMED
11076861
REFERENCE
AUTHORS
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Oheato, N.,
Okazaki, Y., Saito, R., Satoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, P., Takaku-Akai, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
TITLE
Direct Submission
JOURNAL
MEDLINE
20530913
PUBMED
11076861
REFERENCE
AUTHORS
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
COMMENT
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
LOCATION/Qualifiers
1. .3823
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:6430508G12"
/db_xref="taxon:10090"
/clone="6430508G12"
/sex="male"
/tissue_type="olfactory brain"
/clone_lib="RIKEN full-length enriched mouse cDNA library"

misc_feature
275..2706
/dev_stage="adult"
/note="WPI (FRAGMENT) homolog [Homo sapiens]
(SPT)[000307, evidence: FASTV, 95.6%ID, 83.1%length,
match=1698]
putative"
3797..3802
/note="putative"
3823
/note="putative"
polyA_signal
polyA_site
ORIGIN
Query Match 78.1%; Score 487.6; DB 3; Length 3823;
Best Local Similarity 92.2%; Pred. No. 9.7e-133;
Matches 511; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
QY 27 CCCTTTAGTNCAGATGACTTCCCATTCGCGAGGATCTTTGAATGTTGTTGTTGTTGTTA 86
Db 1547 CCTTTAGTACTGATGATCTTCCCATTCGCGAGGATCTTTGAATGTTGTTGTTGTTA 1488
QY 87 TGATCAACAAAGTACCTTACACCTTCCAGAGTATATCTAATTTCCAGACCTTCTGGCAGG 146
Db 1487 TGATCAACAAAGTACCTTACACCTTCCAGAGTATATCTAATTTCCAGACCTTCTGGCAGG 1428
QY 147 GGTTCCTCATTCGTAAGCCTTGAGTCTTGGATCTTCCACCTGGGTGTTGTTGTTGTTA 206
Db 1427 GGTTCCTCATTCGTAAGCCTTGAGTCTTGGATCTTCCACCTGGGTGTTGTTGTTGTTA 1368
QY 207 TGATTCACAAAGTAAACCCCTGCTGTGTAATCCACTCTTTTCCAGACCTGGTGGCAA 266
Db 1367 TGATTCACAAAGTAAACCTGCTGTGTAATCCACTCTTTTCCAGACCTGGTGGTAAG 1308
QY 267 GGTTCCTCATTCGTAAGCCTTGAGTCTTGGATCTTCCACCTGGGTGTTGTTGTTGTTA 326
Db 1307 GGTTCCTCATTCGTAAGCCTTGAGTCTTGGATCTTCCACCTGGGTGTTGTTGTTGTTG 1248
QY 327 AACTGTTGATAGTCTCCCTGCAATTCGTTCCGCTGAGATGCCACTGNTCAAAATTCGG 386
Db 1247 AATTGTTGATAGTCTCCCTGCAATTCGTTCCGCTGAGATGCCACTGNTCAAAATTCGG 1188
QY 387 ACAGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 446
Db 1187 ACAGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1128
QY 447 ACTCTCTTACGATCATCACTCTTCTTCCCACTGGAGGTAAAGGTGTTGTTCTCTCC 506
Db 1127 ACTCTCTTACGATCATCACTCTTCTTCCCACTGGAGGTAAAGGTGTTGTTCTCTCC 1068
QY 507 CATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 566
Db 1067 CATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1008
QY 567 TGTTCCACCCCTG 580
Db 1007 TGTTCCACCCCTG 994
RESULT 6
LOCUS CK232395/c 1113 bp mRNA linear EST 09-DEC-2003
DEFINITION ILLUMINIG_MQ_4036 Katze_MPL2 Macaca mulatta cDNA 5' similar to
human WPI (Hs.191668), mRNA sequence.
ACCESSION CK232395
VERSION CK232395.1 GI:39638753
KEYWORDS EST.
SOURCE Macaca mulatta (rhesus monkey)
ORGANISM Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.
1 (bases 1 to 1113)
REFERENCE Katze, M.G., Thomas, M., Korth, M., Iadonato, S.P. and Magness, C.L.
AUTHORS Large-scale Rhesus Macaque cDNA sequencing
TITLE

JOURNAL
COMMENT

Unpublished (2003)
Contact: C. Magness
Illumigen Biosciences Inc.
2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
Tel: 2063780400
Fax: 2063780408
Email: cmagness@illumigen.com
Sequenced on 2003.12.02. 760 Q20 bases. Assemblies in contig w/ 1 member(s). Contig contains 1 (0%) lib members.
PCR Primers
FORWARD: CCCTCACTAAGGGAACAAA
BACKWARD: CACTATAGCGCAATGGTGA
Insert Length: 1113 Std Error: 0.00
Plate: CL000052 row: C column: 04
Seq primer: CCCTCACTAAGGGAACAAA
POLYA=Yes.

FEATURES
source

1..1113 Location/Qualifiers
/organism="Macaca mulatta"
/mol_type="mRNA"
/strain="Indian"
/db_xref="taxon:9544"
/sex="male"
/dev_stage="newborn infant"
/lab_host="E. coli S04R"
/clone_lib="Katze WMPL2"
/note="Organ: placenta; Vector: Uni-ZAP XR; Site 1: EcoR I; Site 2: Xho I; Created from Stratagene ZAP-CDNA Synthesis kit (catalog #200400) and ZAP-CDNA Gigapack III Gold Cloning Kit (Catalog #200450)."

ORIGIN

Query Match 77.2%; Score 481.8; DB 7; Length 1113;
Best Local Similarity 95.5%; Pred. No. 3.7e-131;
Matches 515; Conservative 0; Mismatches 21; Indels 3; Gaps 2;
QY 27 CCCTTTAGTCAGATGACTCCCATTCGCGAGGATCTTGAATGTTGTTCTTGTTA 86
Db 539 CCTTTAGTTACAGATGACTTCCCATTCGCGAGGATCTTGAATGTTGTTCTTGTTA 480
QY 87 TGATCAACAAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGCCCTCTGGCAGG 146
Db 479 TGATCAACAAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGCCCTCTGGCAGG 420
QY 147 GGTTCCTTCATCTGTGAAGCCTTGAGTCTTTGGATCTTCCACATGGGTTGTTTGTGTTA 206
Db 419 GGTTCCTTCATCTGTGAAGCCTTGAGTCTTTGGATCTTCCCATTTGGGTTGTTTGTGTTG 360
QY 207 TGATTCACAAAGTAAACCTCTGTTGTAATCCACTCTTTTTCCTCC--AGCCCTGGTGGCA 264
Db 359 TGATTCACAAAGTAAACCTCTGTTGTAATCCACTCTTTTCTCCCGCGCTGGTGGCA 300
QY 265 AAGTCCATAGGCTCATTTCTTCGAGCTAACATTTGAAGCCGAATAGAGTATCGTTGTT 324
Db 299 AAGTCCATAGGCTCATTTCTTCGAGCTAACATTTGAAGCCGAATAGAGTATCGTTGTT 240
QY 325 TAACTGTTGCATAGTCCCTGCAATGGTTCGGTTCGATTCGCCACTGNATCAAAATTC 384
Db 239 TAACTGTTGCATAGTCCCTGCAATGGTTCGGTTCGATTCGCCACTGTTTCAAAATTC 180
QY 385 GGACAGATTCATAGTGGCGCTGCCAGTGTCTTCTGCGTGTATGATCCACATAT 444
Db 179 GGACAGATTCATAGTGGCGCTGCCAGTGTCTTCTGCGTGTATGATCCACATAT 120
QY 445 AAACCTCTTACGATCATCAACTCTTTTCCCAACCTGGAGTAAAGTTGTGGTCTCT 504
Db 119 AAACCTCTTACGATCATCAACTCTTTTCCCAACCTGGAGTAAAGTTGTGGTCTCT 60
QY 505 -CCCATGTGTAGTTGAGCAATTTATGACCACATAATAGCTCTACATGAGGATCTTT 562
Db 59 CCCCATGTGTAGTTGAGTATTTATGATCCACATAATAGTGTCTTCAATGAGGATCTTT 1

RESULT 7
CK981350/c

LOCUS
DEFINITION
4113416 BARC 9BOV Bos taurus cDNA clone 9BOV42_007 5', mRNA
sequence.

ACCESSION
VERSION
CK981350
CK981350.1 GI:45499330

KEYWORDS
SOURCE
Bos taurus (cow)
ORGANISM
Bos taurus

REFERENCE
AUTHORS
1 (bases 1 to 654)
Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay, G.P., Bosak, S., Rubenfield, M. and Gasbarre, L.C.
Production of EST from cDNA libraries derived from immunologically activated bovine gut

TITLE
JOURNAL
COMMENT

Unpublished (2004)
Contact: Tad S. Sonstegard
Bovine Functional Genomics Laboratory
Animal and Natural Resources Institute
Bldg. 200 Rm2A BARC-East, Beltsville, MD 20705, USA
Tel: 3015048416
Fax: 3015048414

Email: tads@anri.barc.usda.gov

Single pass sequencing. Bases called and trimmed with phred

0.000925 using options -trim alt -trim fasta. Vector identified

by cross_match using options -mismatch 12 -minscore 12

Plate: 42 row: 0 column: 07

Seq primer: CCCAGTCACGACGTTGTAAACG

High quality sequence stop: 654.

Location/Qualifiers

1..654

/organism="Bos taurus"

/mol_type="mRNA"

/strain="Holstein"

/db_xref="taxon:9913"

/clone="9BOV42_007"

/sex="Male"

/tissue_type="Pooled"

/dev_stage="Multiple"

/lab_host="DH10B T1 phage resistant"

/clone_lib="BARC 9BOV"

/note="Organ: Abomasum; Vector: pagen-1; Site 1: EcoRV;

Site 2: NotI; Equimolar amounts of mRNA extracted from

fundic and pyloric abomasums of 18 and 21 week old steers.

Exposure to Osteragia osteragi was initiated at 15 weeks

of age. fundic and pyloric abomasum"

ORIGIN

Query Match 75.9%; Score 473.4; DB 7; Length 654;
Best Local Similarity 87.6%; Pred. No. 9.9e-129;
Matches 524; Conservative 0; Mismatches 73; Indels 1; Gaps 1;

QY 27 CCCTTTAGTCAGATGACTTCCCATTCGCGAGGATCTTGAATGTTGTTCTTGTTA 86

Db 629 CCCTTTAGTTACAGATGACTTCCCATTCGCGAGGATCTTGAATGTTGTTCTTGTTA 570

QY 87 TGATCAACAAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGCCCTCTGGCAGG 146

Db 569 TGATCAACAAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGCCCTCTGGCAGG 510

QY 147 GGTTCCTTCATCTGTGAAGCCTTGAGTCTTCCACTCTCCACTGGGTTGTTTGTGTTA 206

Db 509 GGTTCCTTCATCTGTGAAGCCTTGAGTCTTCCACTGGGTTGTTTGTGTTA 450

QY 207 TGATTCACAAAGTAAACCTCTGTGTTGAATCCACTCTTTTTCACGCTGGTGGCAAA 266

Db 449 TGATTCACAAAGTAAACCTCTGTGTTGAATCCACTCTTTTTCACGCTGGTGGCAAA 390

QY 267 GGTTCATAGGCTCATTTTCTGCGACTAACATTCGAGCCGAATAGAGTATCGTTGTTA 326


```

QY 548 ACCATGAGGATCTTTCTTTGTTCCACCCCTCATGCGCAANGNTTCTGTGTTGGCATTC 607
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
174 ACCATGAGGATCTTTCTGTGTTCCCA-CCCTGATGCGCAAGCTTCTGTGTTGGTTC 116
QY 608 CAGACTGCTGCCGNACA 624
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
115 CAGACTGTGGCCGACAGA 99

RESULT 9
BE869796/c
LOCUS BE869796
DEFINITION 1101 bp mRNA linear EST 20-OCT-2000
        601446622F1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3850876 5',
        mRNA sequence.
ACCESSION BE869796
VERSION BE869796
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
        Contact: Robert Strausberg, Ph.D.
        Email: cgapbs-remail.nih.gov
        Tissue Procurement: ATCC
        cDNA Library Preparation: Life Technologies, Inc.
        cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
        DNA Sequencing by: Incyte Genomics, Inc.
        Clone distribution: MGC clone distribution information can be
        found through the I.M.A.G.E. Consortium/LLNL at:
        http://image.llnl.gov
        Plate: LLAM9570 row: k column: 05
        High quality sequence stop: 614.
        Location/Qualifiers
            1..1101
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:3850876"
            /tissue_type="adenocarcinoma"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH_MGC_65"
            /notes="Organ: colon; Vector: pCMV-SPORT6; Site 1: NotI;
            Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
            Average insert size 1.8 kb. Library constructed by Life
            Technologies."

FEATURES
source
    1..620
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:3850876"
    /tissue_type="hypothalamus"
    /lab_host="DH10B"
    /clone_lib="NIH_MGC_96"
    /notes="Organ: brain; Vector: pBluescriptR (modified
    pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
    (g-cgg); Oligo-dt primed using primer
    5'-TTTTTTTTTTTTTTVN-3', size-selected for average
    insert size 2.3 kb and normalized to ROT 5. This is a
    primary library enriched for full-length clones and
    constructed using the Cap-trapper method (Carninci, in
    preparation). Library constructed by M. Brownstein
    (NIMH/NHGRI, National Institutes of Health). Note: this is
    a NIH_MGC Library."

ORIGIN
Query Match 68.2%; Score 425.8; DB 4; Length 620;
Best Local Similarity 97.6%; Pred. No. 1.2e-114;
Matches 441; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

```

```

Db 424 CGCTGAGATTGCCACTGTTCAAATTTGGACAGATTTCCATGTTGGCCGCTGCCAGTT 365
QY 417 GTTGTCTCTGTTGTTATGATCCACATAATAAATCTCTCTAGATCATCAACTCTTTTCC 476
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
364 GTTGTCTCTGTTGTTATGATCCACATAATAAATCTCTCTAGATCATCAACTCTTTTCC 305
QY 477 CAACCTGGAGGTAAAGGTTGTGGTCTCTCCATGTGGTAGTTCCAGATATTATGATCCCA 536
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
304 CAACCTGGAGGTAAAGGTTGTGGTCTCTCCATGTGGTAGTTCCAGATATTATGATCCCA 245
QY 537 TAATAGCTCTACCATGAGATCTTTCTTTGTTCCACCCCTGATGGCAANGNTTCTGT 596
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
244 TAATAGCTCTACCATGAGATCTTTCTTTGTTCCCA-CCCTGATGGCAAGGTTTCTGT 186
QY 597 GTTGGCATTCGCCAGACTGCTGCCGNACA 624
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
185 GTTGGCATTCGCCAGACTGCTGCCGTACA 158

```

```

RESULT 10
BI598894/c
LOCUS BI598894
DEFINITION 620 bp mRNA linear EST 07-SEP-2001
        603248001F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5299848 5',
        mRNA sequence.
ACCESSION BI598894
VERSION BI598894
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
        Contact: Robert Strausberg, Ph.D.
        Email: cgapbs-remail.nih.gov
        Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
        cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
        Toshiyuki and Piero Carninci (RIKEN)
        cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
        DNA Sequencing by: Incyte Genomics, Inc.
        Clone distribution: MGC clone distribution information can be
        found through the I.M.A.G.E. Consortium/LLNL at:
        http://image.llnl.gov
        Plate: LLAM11759 row: a column: 01
        High quality sequence start: 7
        High quality sequence stop: 575.
        Location/Qualifiers
            1..620
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:5299848"
            /tissue_type="hypothalamus"
            /lab_host="DH10B"
            /clone_lib="NIH_MGC_96"
            /notes="Organ: brain; Vector: pBluescriptR (modified
            pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
            (g-cgg); Oligo-dt primed using primer
            5'-TTTTTTTTTTTTTTVN-3', size-selected for average
            insert size 2.3 kb and normalized to ROT 5. This is a
            primary library enriched for full-length clones and
            constructed using the Cap-trapper method (Carninci, in
            preparation). Library constructed by M. Brownstein
            (NIMH/NHGRI, National Institutes of Health). Note: this is
            a NIH_MGC Library."

```

FEATURES

source

```

1..620
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5299848"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/clone_lib="NIH_MGC_96"
/notes="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
(g-cgg); Oligo-dt primed using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average
insert size 2.3 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIMH/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC Library."

```

ORIGIN

```

Query Match 68.2%; Score 425.8; DB 4; Length 620;
Best Local Similarity 97.6%; Pred. No. 1.2e-114;
Matches 441; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

```

```
QY 27 CCCTTTAGTNCAGATGACTCCCATTCGCGAGGATCTTTGAATGTTGTTGTTCTTGTGTTA 86
Db 460 CCTTTAGTTACAGATGACTCCCATTCGCGAGGATCTTTGAATGTTGTTGTTCTTGTGTTA 401
QY 87 TGATCAACAAGTACCTTACACCTTTCAGAGTATATCTAATT-TCCCAGCCTTCTGGCAG 145
Db 400 TGATCAACAAGTACCTTACACCTTTCAGAGTATATCTAATTGTTCCAGCCTTCTGGCAG 341
QY 146 GGGTCTTCTATTCTGTAAAGCCTTGAGTCTTTGGATCTTCCACCTGGGTGTTGTTGTT 205
Db 340 GGGTCTTCTATTCTGTAAAGCCTTGAGTCTTTGGATCTTCCACCTGGGTGTTGTTGTT 281
QY 206 ATGATTCACAAGTAAACCCCTGCTGTGATCACTCTTTTTCAGCCTGGTGGCAA 265
Db 280 ATGATTCACAAGTAAACCCCTGCTGTGATCACTCTTTTTCAGCCTGGTGGCAA 221
QY 266 AGGTCCATAAGGGTCAATTTCTGCAGCTAAACATTAAGAGCGAATAGAGGTATCGTTGGTT 325
Db 220 AGGTCCATAAGGGTCAATTTCTGCAGCTAAACATTAAGAGCGAATAGAGGTATCGTTGGTT 161
QY 326 AAACCTGTTGCATAGCTCCCTGCAATGGTTCGCTGAGATGGCCACTGNTCAAAATTNCG 385
Db 160 AAACCTGTTGCATAGCTCCCTGCAATGGTTCGCTGAGATGGCCACTGNTCAAAATTNCG 101
QY 386 GACAGATTCATGATGGCGCTGCCAGTGTGTTGTTCTGTTGTTATGATCCACATATA 445
Db 100 GACAGATTCATGATGGCGCTGCCAGTGTGTTGTTCTGTTGTTATGATCCACATATA 41
QY 446 AACTCTCTACGATCATCACTCTTCTTCCC 477
Db 40 AACTCTCTACGATCATCACTCTTCTTCCC 9
```

```
RESULT 11
BU509208/c
LOCUS BU509208 891 bp mRNA linear EST 12-SEP-2002
DEFINITION AGENCOURT_10099965 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6503763
5', mRNA sequence.
ACCESSION BU509208
VERSION BU509208.1 GI:22815441
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 891)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14061 row: p column: 04
High quality sequence stop: 649.
Location/Qualifiers
1. .891
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6503763"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site: 1; NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb."
```

FEATURES
source

ORIGIN

```
Query Match 66.7%; Score 416; DB 5; Length 891;
Best Local Similarity 98.1%; Pred. No. 1.1e-111;
Matches 419; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 27 CCCTTTAGTNCAGATGACTCCCATTCGCGAGGATCTTTGAATGTTGTTGTTCTTGTGTTA 86
Db 427 CCTTTAGTTACAGATGACTCCCATTCGCGAGGATCTTTGAATGTTGTTGTTCTTGTGTTA 368
QY 87 TGATCAACAAGTACCTTACACCTTTCAGAGTATATCTAATT-TCCCAGCCTTCTGGCAGG 146
Db 367 TGATCAACAAGTACCTTACACCTTTCAGAGTATATCTAATT-TCCCAGCCTTCTGGCAGG 308
QY 147 GGTCTTCTATTCTGTAAAGCCTTGAGTCTTTGGATCTTCCACCTGGGTGTTGTTGTTA 206
Db 307 GGTCTTCTATTCTGTAAAGCCTTGAGTCTTTGGATCTTCCACCTGGGTGTTGTTGTTA 248
QY 207 TGATTCACAAGTAAACCCCTGCTGTGATCACTCTTTTTCAGCCTGGTGGCAA 266
Db 247 TGATTCACAAGTAAACCCCTGCTGTGATCACTCTTTTTCAGCCTGGTGGCAA 188
QY 267 GGTCCATAAGGGTCAATTTCTGCAGCTAAACATTAAGAGCGAATAGAGGTATCGTTGTTA 326
Db 187 GGTCCATAAGGGTCAATTTCTGCAGCTAAACATTAAGAGCGAATAGAGGTATCGTTGTTA 128
QY 327 AACTGTTGCATAGCTCCCTGCAATGGTTCGCTGAGATGGCCACTGNTCAAAATTNCGG 386
Db 127 AACTGTTGCATAGCTCCCTGCAATGGTTCGCTGAGATGGCCACTGNTCAAAATTNCGG 68
QY 387 ACAGATTCATGATGGCGCTGCCAGTGTGTTGTTCTGTTGTTATGATCCACATATAA 446
Db 67 ACAGATTCATGATGGCGCTGCCAGTGTGTTGTTCTGTTGTTATGATCCACATATAA 8
QY 447 ACTCTCTC 453
Db 7 ACTCTTC 1
RESULT 12
AL049041/c
LOCUS AL049041 630 bp mRNA linear EST 04-SEP-2003
DEFINITION DKFZp434P0218_r1.434 (synonym: hte83) Homo sapiens cDNA clone
DKFZp434P0218, mRNA sequence.
ACCESSION AL049041
VERSION AL049041.1 GI:4728350
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 630)
AUTHORS Ottenwaelder, B., Obermaier, B., Mewes, H.W., Gassenhuber, J. and
Wiemann, S.
TITLE EST (Ottenwaelder, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS
MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by MediGenomix (Martinsried/Germany) within the cDNA
sequencing consortium of the German Genome Project. No s1 sequence
available.
This clone (DKFZp434P0218) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1. .630
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp434P0218"
```


Plate: 06 row: c column: 12
Seq primer: -21M13 forward primer (ABI).
Location/Qualifiers

FEATURES

source

1. .474
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NHTBC cn06c12"
/sex="Female"
/tissue_type="Bone"
/cell_type="Trabecular Bone Cells"
/lab_host="SURE"
/clone_lib="Normal Human Trabecular Bone Cells"
/notes="Organ: Hip; Vector: pBluescript; Site: 1; EcoRI;
Library constructed by Dr. Marian Young and Dr. Pamela
Gehron Robey (NIDCR)"

ORIGIN

Query Match 60.3%; Score 376; DB 1; Length 474;
Best Local Similarity 97.0%; Pred. No. 6.3e-100;
Matches 390; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

QY 223 CCCTGCTGTGTGAATCCACATCTTTTCCAGAGCTGTGGCAAGGTCCATAAGGGTTCAT 282
DB 474 CCTGTCTGTGTGAATCCACATCTTTTCCAGAGCTGTGGCAAGGTCCATAAGGGTTCAT 415

QY 283 TTTCTGCAGCTAACATTGAAGCCGAATAGAGGTATCGTTGGTTAAACTGTTGCATAGCTC 342
DB 414 TTTCTGCAGCTAACATTGAAGCCGAATAGAGGTATCGTTGGTTAAACTGTTGCATAGCTC 355

QY 343 CTGCAATTGGTCCGCTGAGATTGCCACTGNTCAAAATTCGACAGATTCATCGATG 402
DB 354 CTGCAATTGGTCCGCTGAGATTGCCACTGNTCAAAATTCGACAGATTCATCGATG 295

QY 403 GCCGCTGCCAGCTGTGTCTGTGTGTATGATCCACATAATAAATTCCTCTAGCATCAT 462
DB 294 GCCGCTGCCAGCTGTGTCTGTGTGTATGATCCACATAATAAATTCCTCTAGCATCAT 235

QY 463 CAACTCTTCTTCCCACTGGAGGTAAAGGTGTGTCTCTCCCATGTGTAGTTCGAG 522
DB 234 CAACTCTTCTTCCCACTGGAGGTAAAGGTGTGTCTCTCCCATGTGTAGTTCGAG 175

QY 523 CATTATGANCACATAATACGCTTACCATGAGGATCTTTCTTTTGTGTCCACCCCTGAT 582
DB 174 TATTATGATCCACATAATAGGTCTTACCATGAGGATCTTTCTTTTGTGTCCCA-CCCTGAT 116

QY 583 GGCAGAGTTCGTGTGTGGCATTCACAGACTGTGCCGNACA 624
DB 115 GGCAGAGTTCGTGTGTGGCATTCACAGACTGTGCCGTACA 74

RESULT 15

BF369480
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BF369480 557 bp mRNA linear EST 24-NOV-2000
RC0-GN0090-260900-032-e09 GN0090 Homo sapiens cDNA, mRNA sequence.
BF369480
GI:11331505

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 557)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsumura, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

PUBMED

COMMENT

10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC0&t2=RC0-GN0090-
260900-032-e09&t3=2000-09-28&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 16
High quality sequence stop: 557.

FEATURES

source

1. 557
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="GN0090"
/note="Organ: Placenta normal; Vector: puc18; Site: 1;
SmaI; Site: 2; SmaI; A mini-library was made by cloning
products derived from ORBESTES PCR (U.S. Letters Patent
application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the pUC 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."

ORIGIN

Query Match 59.9%; Score 373.6; DB 2; Length 557;
Best Local Similarity 97.7%; Pred. No. 3.4e-99;
Matches 388; Conservative 0; Mismatches 7; Indels 2; Gaps 1;

QY 27 CCCTTAGTNCAGATGACTTCCCATTCGCGAGGATCTTGAATGTTGTTCTGTGTTA 86
DB 161 CCTTAAATACAGATGACTTCCCATTCGCGAGGATCTTGAATGTTGTTCTGTGTTA 220

QY 87 TGATCAACAAAGTACCTTACACCTTCACAGTATATCTAATTTCCACAGCTTCTGGCAGG 146
DB 221 TGATCAACAAAGTACCTTACACCTTCGCGAGTATATCTAATTTCCACAGCTTCTGGCAGG 280

QY 147 GGTTCCTTCTGTGTAAAGCCTTGAGTTCCTTGGATCTTCCACCTGGGTGTTTGTGTTA 206
DB 281 GGTTCCTTCTGTAAAGCCTTGAGTTCCTTGGATCTTCCACCTGGGTGTTTGTGTTA 340

QY 207 TGATTCACAAAGTAAACCTGTCTGTGAATCCACTC--TTTTTCCACGCTGTGGCA 264
DB 341 TGATTCACAAAGTAAACCTGTCTGTGAATCCACTCTTTTTTCCACGCTGTGGCA 400

QY 265 AAGTCCATAAGGTCATTTTCTGAGCTAACCAATTGAAGCCGAATAGAGGTATCGTTGGT 324
DB 401 AAGTCCATAAGGTCATTTTCTGAGCTAACCAATTGAAGCCGAATAGAGGTATCGTTGGT 460

QY 325 TAAACTGTTGCATAGCTCCCTGCAATTTGGTTCGCTGAGATTGCCACTGNTCAAAATTC 384
DB 461 TAAACTGTTGCATAGCTCCCTGCAATTTGGTTCGCTGAGATTGCCACTGNTCAAAATTC 520

QY 385 GGACAGATTCCATGGTAGCCGCTGCCACGTTGTTGT 421
DB 521 GGACAGATTCCATGGTAGCCGCTGCCACGTTGTTGT 557

Search completed: June 30, 2005, 06:29:09

Job time : 2875 secs

THIS PAGE BLANK (USPTO)